第一章

不分明子集论初步

本章的主要目在于明确不分明子集的定义以及为处理不分明概率所必要的与不分明集合有关的事项。在这里，把不分明集看成经典集合的特征函数的一般化，并从这个观点来叙述不分明集合和它与经典集合间的联系。

这一章是预备性的，但却是重要的。如果读者已学过不分明数学或类似的内容，则本章可以不读而直接阅读第二章。

§1 经典集合

在数学中，我们常常需要去研究具有某些特定性质的对象的总体。这样定义了的任何对象的总体，我们称它为集合，每一个属于这种集合的对象，则称为集合的一个元素。

集合的元素可以是任意种类的对象：点，数，函数，事件，人等等。例如(1) 一切有理数组成的集合；(2) 在给定平面上全体的点组成的集合；(3) 1984年毕业的大学生组成的集合；(4) 西安市三路公共汽车全体组成的集合。

集合有这么一个特性，对于一个指定的集合和一个给定的元素，我们可以判定该元素属于这个集合，还是不属于这个集合。二者必居其一，决不能有第三种情况。例如，对于给定的一个数$\sqrt{2}$，因为它是无理数，所以它不是由(1)所给定集合的元素。

在实际工作中，为了研究问题的方便，总是把议题局限在某一个范围内，这个范围我们常称其为“论域”.显然论域也是一个集合。论域我们常用大写字母 X，U，Ω，Y 等来表示。而其中元素常用小写字母 a，b，x，y，…来表示。

给定一个论域 X，X 中某一部分元素的全体，叫做 X 的一个子集或 X 中的一个集合。换句话说，假如有两个集合 X 和 A，A 中的每一个元素属于 X，则称 A 是 X 的一个子集。且记为

$$A \subset X \text{ 或 } X \supset A$$

当 A 仅含一个元素 x 时，我们用记号 $x \in X$ 来表示 x 属于 X。而元素 y 不属于 X 时，我们记作

$$y \overline{\in} X \text{ (或 } y \notin X\text{)。}$$

以 $\varnothing$ 表示空集，X 表示全集(论域)。

任给一性质或条件 P，$P(x)$ 表示“x 具有性质(条件)P”，则

$$A = \{x | P(x)\}$$

表示 X 中具有性质 P 的一切元素构成的子集。以“$\forall x$，$P(x)$”表示对所有 x 均有性质 P，“$\exists x$，$P(x)$”表示存在 x 具有性质 P。这种记法，当 x 为集合时仍然使用，如 $\mathscr{P}(X) = \{A | A \subseteq X\}$ 为 X 的一切子集的集合，称 $\mathscr{P}(X)$ 为 X 的幂集，约定 $\varnothing$，$X \in \mathscr{P}(X)$。如果 X 中有 n 个元素，则 $\mathscr{P}(X)$ 中有 2^n 个元素。例如，设 $X = \{a, b, c\}$，则 $\mathscr{P}(X) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\}$。

例 1.1.1 设 $A = \{x | x^2 - 1 = 0\}$，它表示满足方程 $x^2 - 1 = 0$ 的一切 x 组成的集合。解方程即得

$$A = \{x | x^2 - 1 = 0\} = \{-1, 1\},$$

它的元素只有两个：-1和1。

例 1.1.2 设 Z 是整数集，那么

$$E=\left\{n\left|\frac{n}{2}\in Z\right.\right\}$$

也是一个集合。条件“$\frac{n}{2}\in Z$”表示$\frac{n}{2}$属于整数集，即这种 n 必须是整数：

$$E=\left\{n\left|\frac{n}{2}\in Z\right.\right\}=\{0,\pm 2,\pm 4,\pm 6,\cdots\}.$$

设 A、B 是 X 中的两个子集，若 $x\in A$ 时必有 $x\in B$，称 A 含于 B，或 B 包含 A，记作 $A\subset B$。显然，包含关系具有以下性质：

(1)（自反性）$A\subset A$；

(2)（对称性）若 $A\subset B$，$B\subset A$，则 $A=B$；

(3)（传递性）若 $A\subset B$，$B\subset C$，则 $A\subset C$。

设 A、$B\in\mathscr{P}(X)$。$A\cup B$、$A\cap B$ 和 A^c（或 $\bar{A}$）分别称为 A 与 B 的并集、交集和 A 的补集。其定义如下：

$$A\cup B=\{x\in X\mid x\in A \text{ 或 } x\in B\};$$

$$A\cap B=\{x\in X\mid x\in A \text{ 且 } x\in B\};$$

$$A^c=\{x\in X\mid x\,\overline{\in}\, A\}.$$

我们知道，如果 X 的幂集$\mathscr{P}(X)$关于并与交运算满足以下诸条件，则称 $\mathscr{P}(X)$ 构成一σ代数，记为（$\mathscr{P}(X)$，$\cup$，$\cap$，C）：

1）$X\in\mathscr{P}(X)$；

2）若$A_i\in\mathscr{P}(X)$，$i=1,2,\cdots$，则$\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{P}(X)$；

3）若 $A\in\mathscr{P}(X)$，则 $A^c\in\mathscr{P}(X)$。

我们上面给出的 $\mathscr{P}(X)$，显然为一 σ 代数。

不难验证，$(\mathscr{P}(X), \cup, \cap, C)$ 具有以下性质：

(1) $A\cap B=B\cap A$，$A\cup B=B\cup A$ （交换律）

(2) $(A\cup B)\cup C=A\cup(B\cup C)$

$(A\cap B)\cap C=A\cap(B\cap C)$ （结合律）

(3) $A\cup\varnothing=A$，$A\cap X=A$ （单位元存在性）

(4) $A\cup A^c=X$，$A\cap A^c=\varnothing$ （互补律）

(5) $A\cup(A\cap B)=A$，$A\cap(A\cup B)=A$ （吸收律）

(6) $A\cup(B\cap C)=(A\cup B)\cap(B\cup C)$

$A\cap(B\cup C)=(A\cap B)\cup(B\cap C)$ （分配律）

(7) $A\cup A=A$，$A\cap A=A$ （幂等律）

(8) $A\cup X=X$，$A\cap\varnothing=\varnothing$， （两极律）

(9) $(A^c)^c=A$ （对合律）

(10) $(A\cup B)^c=A^c\cap B^c$，$(A\cap B)^c=A^c\cup B^c$。

（对偶律）

性质(1)—(10)可扩展至 $B_t\in\mathscr{P}(X)(t\in T)$ 的情形。

记号

$$f: A\to B$$

表示 A 到 B 的映射，即对于任给的 $x\in A$，有 $y=f(x)$ 与之对应。A 称为 f 的定义域，而 $f(A)=\{y\mid\exists x\in A, y=f(x)\}$ 称为 f 的值域。

如果 $A\in\mathscr{P}(X)$，称

$$I_A: X\to\{0,1\}$$

为 A 的特征函数，或简称为 A 的指示子，用

$$I_A(x)=\begin{cases}1; & x\in A,\\0; & x\notin A.\end{cases}$$

来表示。一切特征函数的集合记为

$$\mathscr{F}_0(X)=\{I_A|I_A: X\to\{0,1\}\}.$$

若在$\mathscr{F}_0(X)$中引入运算$\vee$，$\wedge$与$'$如下：

$$I_A(x)\vee I_B(x)=\max(I_A(x),\ I_B(x)),$$
$$I_A(x)\wedge I_B(x)=\min(I_A(x),\ I_B(x)),$$
$$I'_A(x)=1-I_A(x).$$

显然，这是布尔和、积与补的运算：

$\vee$	0	1
0	0	1
1	1	1

$\wedge$	0	1
0	0	0
1	0	1

A	A^c
0	1
1	0

例1.1.3 设 $X=\{a,b,c,d,e\}$，$A=\{b,c,e\}$。$B=\{a,d,e\}$。于是有

$$I_A(a)=0,\ I_A(b)=1,\ I_A(c)=1,\ I_A(d)=0,$$
$$I_A(e)=1.$$
$$I_B(a)=1,\ I_B(b)=0,\ I_B(c)=0,\ I_B(d)=1,$$
$$I_B(e)=1.$$

从而得

$$I_A(a)\vee I_A(a)=\max(0,1)=1,$$
$$I_A(a)\wedge I_A(a)=\min(0,1)=0,$$
$$I'_A(a)=1-I_A(a)=1-0=1.$$

	a	b	c	d	e
$I_A(\cdot)\vee I_B(\cdot)$	1	1	1	1	1
$I_A(\cdot)\wedge I_B(\cdot)$	0	0	0	0	1
$I'_A(\cdot)$	1	0	0	1	0

众所周知，$\mathscr{P}(X)$和$\mathscr{F}_0(X)$作为集合是等同的，即$\mathscr{P}(X)$

$\cong \mathscr{F}_0(X)$。事实上，我们可以定义两个映射

f: $\mathscr{P}(X) \to \mathscr{F}_0(X)$，即 $f(A) = I_A$；

g: $\mathscr{F}_0(X) \to \mathscr{P}(X)$，即 $g(I_A) = \{x \in X \mid I_A(x) = 1\}$.

使

$$f \circ g = f(g(I_A)) = f(\{x \in X \mid I_A = 1\}) = I_A,$$

$$g \circ f = g(f(A)) = g(I_A) = A.$$

这说明在 $\mathscr{P}(X)$ 和 $\mathscr{F}_0(X)$之间存在一一对应的映射，按照等同定义，集合 $\mathscr{P}(X)$和 $\mathscr{F}_0(X)$等同。等同的集可 以认为是相同的，$\mathscr{P}(X)$ 看作是直观的图式模型，而 $\mathscr{F}_0(X)$作 为函数空间的数学模型。用 $\mathscr{F}_0(X)$ 不如用 $\mathscr{P}(X)$ 直观易懂，但可进行更为一般的讨论，由它可引入不分明集合的概念。

易证 $(\mathscr{P}(X), \cup, \cap, C) \cong \mathscr{F}_0(X), \vee, \wedge, ')$，

即 $\mathscr{P}(X) \cong \mathscr{F}_0(X)$；$I_{A\cup B} = I_A \vee I_B$；

$I_{A\cap B} = I_A \wedge I_B$；$I_{A^c} = I'_A$.

例 1.1.4 设论域为 $X = \{a,b\}$，则

$$\mathscr{P}(X) = \{\varnothing, \{a\}, \{b\}, X\}.$$

由于

1) $X \in \mathscr{P}(X)$，$\varnothing \in \mathscr{P}(X)$；

2) $\{a\} \in \mathscr{P}(X) \Rightarrow \{a\}^c = \{b\} \in \mathscr{P}(X)$，

$\{b\} \in \mathscr{P}(X) \Rightarrow \{b\}^c = \{a\} \in \mathscr{P}(X)$；

3) $\{a\}$，$\{b\} \in \mathscr{P}(X)$，$\{a\} \cap \{b\} = \varnothing$ 有

$\{a\} \cup \{b\} = \{a,b\} = X \in \mathscr{P}(X)$.

所以 $\mathscr{P}(X)$是一布尔代数．又设 I_A: $\mathscr{P}(X) \to \{0,1\}$，即 $\varnothing \to 0$，$X \to 1$，$\{a\} \to I_{\{a\}}$，$\{b\} \to I_{\{b\}}$，故$\mathscr{P}(X) \to \mathscr{F}_0(X) = \{0, I_{\{a\}}, I_{\{b\}}, 1\}$。在$\mathscr{F}_0(X)$有

1′) $0 \in \mathscr{F}_0(X)$，$1 \in \mathscr{F}_0(X)$，

2′) $I_{\{a\}} \in \mathscr{F}_0(X) \Rightarrow I'_{\{a\}} = I_{\{b\}} \in \mathscr{F}_0(X)$,

$I_{\{b\}} \in \mathscr{F}_0(X) \Rightarrow I'_{\{b\}} = I_{\{a\}} \in \mathscr{F}_0(X)$,

3′) $I_{\{a\}}$, $I_{\{b\}} \in \mathscr{F}_0(X) \Rightarrow$

$$I_{\{a\}\cup\{b\}} = I_{\{a\}} \vee I_{\{b\}} \in \mathscr{F}_0(X)$$

从而$(\mathscr{F}_0(X), \vee, \wedge, ')$为一布尔代数。且$(\mathscr{P}(X), \cup, \cap, c)$和$(\mathscr{F}_0(X), \vee, \wedge, ')$等同.因此将$\mathscr{P}(X)$和$\mathscr{F}_0(X)$看作相同的(上面所述之$f$为$I_A$，$g$为$I_A$的逆映射$I_A^{-1}$)。于是关于$\mathscr{P}(X)$上的讨论将移至$\mathscr{F}_0(X)$上讨论。

下面我们来看一下在$\mathscr{F}_0(X)$上如何表示一个集合。

例 1.1.5 设在例 1.1.1 的条件下，集合A记为

$$A = \{(a,0), (b,1), (c,1), (d,1), (e,0)\},$$

或写成

$$A = 0/a + 1/b + 1/c + 1/d + 0/e$$
$$= 1/b + 1/c + 1/d.$$

而

$$B = \{(a,1), (b,0), (c,0), (d,1), (e,1)\}$$
$$= 1/a + 1/d + 1/e.$$

因此关于集合的运算结果有

$$A \cup B = \{(a,1), (b,1), (c,1), (d,1), (e,1)\}$$
$$A \cap B = \{(a,0), (b,0), (c,0), (d,1), (e,0)\}$$
$$A^c = \{(a,1), (b,0), (c,0), (d,0), (e,1)\},$$

或写成

$$A \cup B = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} + \frac{1}{e};$$

$$A \cap B = \frac{1}{d},$$

$$A^c = \frac{1}{a} + \frac{1}{e}.$$

而从可见，在($\mathscr{F}_0(X)$，∨，∧；0，1)中集的运算，实际上是按逐个元素对其特征函数值进行运算，这里“+”仅是一种符号，表示“总括”之意。因此，一个集合 A 一般可表示为

$$A=\{(x,I_A(x)),\ x\in X\}$$

或

$$A=\sum_i I_A(x_i))/x_i,\ x_i\in X.$$

另外，把特征函数 $I_A(x)$ 理解为 x 隶属于 A 的程度，x 属于 A 的程度为1，即 $I_A(x)=1$，则 x 属于 A。从这个观点出发，扩展特征函数的工作将会顺利得多。

§2 F 集 合

设 X 是经典集合，为论域。

定义 1.2.1 (Zadeh,1965)映射 $\underset{\sim}{A}: X\to[0,1]$ 称为不分明集合(Fuzzy Sets)，简称为 F 集。$\underset{\sim}{A}(x)$ 称为 x 相对于 F 集合 $\underset{\sim}{A}$ 的隶属程度。$\underset{\sim}{A}(\cdot)$ 称为 F 集合 $\underset{\sim}{A}$ 的隶属函数.

如果所讨论的集合是经典的，那么 $A(x)$ 是 A 的特征函数 $I_A(x)$，亦是 A 的隶属函数。若 $A(x)=1$，则 x 完全属于集合 A，即 $x\in A$。若 $A(x)=0$，则 x 完全不属于集合 A，即 $x\notin A$。x 属于 A 或 x 不属于 A 是完全确定的。但是，对于反映不分明(模糊)概念的 F 集合 $\underset{\sim}{A}$ 来说，$\underset{\sim}{A}(x)$ 表示 x 属于 F 集合 $\underset{\sim}{A}$ 的程度。它可取0与1间的任何一个数值。如 $\underset{\sim}{A}(x)=0.7$，表示 x 属于 $\underset{\sim}{A}$ 的程度为0.7，若 $\underset{\sim}{A}(x_1)=0.4$，就说 x 比 x_1 相对的更属于 F 集合 $\underset{\sim}{A}$。

可见，隶属函数的概念是特征函数概念的一般化。从而用特征函数刻划的经典集合是用隶属函数刻划的 F 集合的特

殊情形，F 集合是经典集合的一般化。

例 1.2.1 设论域为 $X=\{a,b,c,d,e\}$，它们是五个不同苹果的集合。我们关心的是“烂苹果”$\underset{\sim}{A}$。按照经典集合的观点，每个苹果要么属于 $\underset{\sim}{A}$ 要么不属于 $\underset{\sim}{A}$。即 $I_A(x)$要么为1要么为 0，别无其它选择。可仔细想一下实际情形，可能有些苹果，把它算作烂苹果，并不完全恰当，把它作为好苹果却有点勉强。换句话说，该苹果属于 $\underset{\sim}{A}$ 的程度不是 1 也不是 0，而是 0 与 1 中间的一个值。这类不分明的概念(无明确外延的概念)就是引进 F 集合的实际背景。

对于 X 中的每个 x 伴随着一个隶属函数 $\underset{\sim}{A}(x)$，因此一般情况下一个 F 集合可表示为

$$\underset{\sim}{A}=\{(x,\underset{\sim}{A}(x))\mid x\in X\},$$

如果 X 是有限集或可数集，可表示为

$$\underset{\sim}{A}=\sum \underset{\sim}{A}(x_i)/x_i;$$

如果 X 是无限不可数集，可表示为

$$\underset{\sim}{A}=\int \underset{\sim}{A}(x)/x.$$

例 1.2.2 某单位招收服务人员。现需对五名应征者 x_1，x_2，x_3，x_4，x_5 进行评议。此时

$$x=\{x_1,x_2,x_3,x_4,x_5\}$$

评议内容是他们的仪表美。按百分制打分为

x_1：85分； x_2：75分； x_3：98分；

x_4：30分； x_5：60分。

若将它们都除以 100 分，便给出一个从 X 到[0,1]闭区间上的映射 $\underset{\sim}{A}: x\to[0,1]$。即对每个 x_i 给出了隶属度 $\underset{\sim}{A}(x_i)$：

$$\underset{\sim}{A}(x_1)=0.85;\ \underset{\sim}{A}(x_2)=0.75;\ \underset{\sim}{A}(x_3)=0.98;$$

$\underset{\sim}{A}(x_4)=0.30$; $\underset{\sim}{A}(x_5)=0.60$.

这样就确定了一个 F 集合 $\underset{\sim}{A}$，表示这五名应征者对“仪表美”这个不分明概念的符合程度。记为

$$\underset{\sim}{A}=\{0.85,0.75,0.98,0.30,0.60\},$$

或

$$\underset{\sim}{A}=\{(x_1,0.85),(x_2,0.75),(x_3,0.98),(x_4,0.30),(x_5,0.60)\},$$

或

$$\underset{\sim}{A}=0.85/x_1+0.75/x_2+0.98/x_3+0.30/x_4+0.60/x_5.$$

顺便提及，F 集合表达方式的选择可由具体问题而定。这最后一种记法是 Zadeh 引进的，它仅是一种记法，不是分式的求和。其“分母”是论域 X 的元素，“分子”是相应元素的隶属度，“+”表示由哪些项构成，仍是总括之意，并非“相加”。今后我们约定，隶属度为 0 的项可以不写出。如

$$\underset{\sim}{A}=1/a+0.8/b+0/c+0.2/d$$

和

$$\underset{\sim}{A}=1/a+0.8/b+0.2/d$$

表示同一个 F 子集合。显然，当 X 的元素是不可数无限集时，和号用“$\int$”代替，它不是积分，因为其后无 dx。

例 1.2.3 以年龄为论域，取 $X=[0,100]$。Zadeh给出“年老($\underset{\sim}{A}$)”和“年轻($\underset{\sim}{B}$)”两个 F 集合的隶属函数为

$$\underset{\sim}{A}(x)=\begin{cases}0, & 0\leqslant x\leqslant 50,\\ \left[1+\left(\dfrac{x-50}{5}\right)^{-2}\right]^{-1}, & 50<x\leqslant 100.\end{cases}$$

$$\underset{\sim}{B}(x)=\begin{cases}1, & 0\leqslant x\leqslant 25,\\ \left[1+\left(\dfrac{x-25}{5}\right)^{2}\right]^{-1} & 25<x\leqslant 100,\end{cases}$$

用图表示如下：

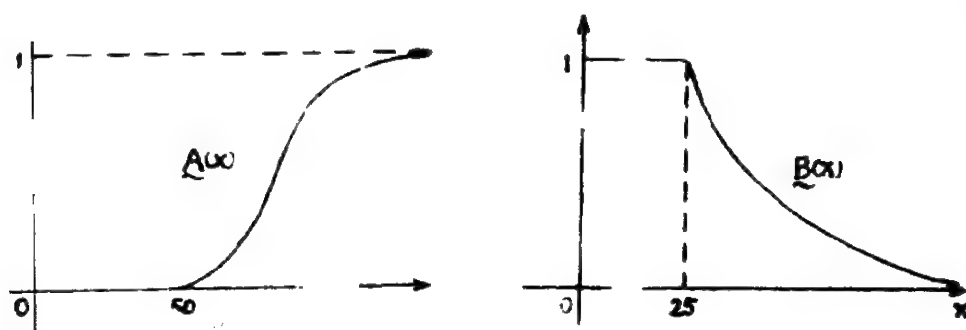


图 1.2.1

如果 $x=60$，则 $\underset{\sim}{A}(60)=0.8$，$\underset{\sim}{B}(60)=0.2$，故可认为60岁是比较年老的。

例 1.2.4 运用电子计算机自动识别几何图形，识别对象为三角形，取论域为：

$$X=\{(A,B,C)\mid A+B+C=180°,A\geqslant B\geqslant C\geqslant 0\}$$

即把三角形三内角的角度从大到小所可能出现的每一种排列，当作 X 的一个元素。用 $\underset{\sim}{I}$ 表示"近似等腰三角形"，$\underset{\sim}{R}$ 表示"近似直角三角形"，$\underset{\sim}{E}$ 表示"近似正三角形"，其隶属函数分别规定如下：

$$\underset{\sim}{I}(A,B,C)=1-\frac{1}{60}\min(A-B,B-C),$$

$$\underset{\sim}{R}(A,B,C)=1-\frac{1}{90}|A-90|,$$

$$\underset{\sim}{E}(A,B,C)=1-\frac{1}{180}(A-C).$$

容易验证，当 $A=B$ 或 $B=C$ 时，$\underset{\sim}{I}(A,B,C)=1$，$\underset{\sim}{I}(120,60,0)=0$；当 $A=90°$ 时，$\underset{\sim}{R}(A,B,C)=1$，$\underset{\sim}{R}(180,0,0)=0$；当 $A=B=C$ 时，$\underset{\sim}{E}(60,60,60)=1$，$\underset{\sim}{E}(180,0,0)=0$。现有一个

三角形，测得三内角之度数为(80,55,45)，试问它应该算作哪一类型的三角形？

由于算得 $\underset{\sim}{I}(80,55,45)=\frac{5}{6}$ ；$\underset{\sim}{R}(80,55,45)=\frac{8}{9}$；$\underset{\sim}{E}(80,55,45)=1-\frac{35}{180}=0.81$。其中 $\underset{\sim}{R}(80,55,45)$最大，故可以将它归入近似直角三角形。

读者至此不难看出，一切隶属函数的全体

$$\mathscr{F}(X)=\{\underset{\sim}{A}(\cdot)|\underset{\sim}{A}: X\to[0,1]\}$$

与 X 上的一切 F 集合的全体 $\mathscr{P}_F(X)$是等同的。后者与经典集合之$\mathscr{P}(X)$相应。

容易给出 F 集合的几何表示。考虑用 $\underset{\sim}{A}(x)$作纵坐标，从而每个元素 x 的隶属度由它的纵坐标表示。横坐标为x，一切 x 构成 X。而阴影部分表示 F 集合 $\underset{\sim}{A}$，如图 1.2.2所示。

仿照$\mathscr{F}_0(X)$上给出的集合运算关系可给出 F 集合之间的各种运算定义如下：

定义 1.2.2 设 $\underset{\sim}{A}$，$\underset{\sim}{B}\in\mathscr{F}(X)$。若 $\forall x\in X$，有 $\underset{\sim}{A}(x)\leqslant\underset{\sim}{B}(x)$，称 $\underset{\sim}{A}$ 含于 $\underset{\sim}{B}$，或 $\underset{\sim}{B}$ 包含 $\underset{\sim}{A}$，并记作 $\underset{\sim}{A}\subset\underset{\sim}{B}$。若 $\forall x\in X$，$\underset{\sim}{A}(x)=\underset{\sim}{B}(x)$，称 $\underset{\sim}{A}$ 等于 $\underset{\sim}{B}$，记作 $\underset{\sim}{A}=\underset{\sim}{B}$。

ϕ表示隶属函数恒为 0 的 F 集合，X 表示隶属函数恒为 1 的 F 集合。

定义 1.2.3 设 $\underset{\sim}{A}$、$\underset{\sim}{B}\in\mathscr{F}(X)$，$\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 的并 $\underset{\sim}{A}\cup\underset{\sim}{B}$，交 $\underset{\sim}{A}\cap\underset{\sim}{B}$，$\underset{\sim}{A}$ 的补集 $\underset{\sim}{A}^C$ 的隶属函数分别为

$$(\underset{\sim}{A}\cup\underset{\sim}{B})(x)=\underset{\sim}{A}(x)\vee\underset{\sim}{B}(x)=\max(\underset{\sim}{A}(x),\underset{\sim}{B}(x)),$$

$$(\underset{\sim}{A}\cap\underset{\sim}{B})(x)=\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x)=\min(\underset{\sim}{A}(x),\underset{\sim}{B}(x)),$$

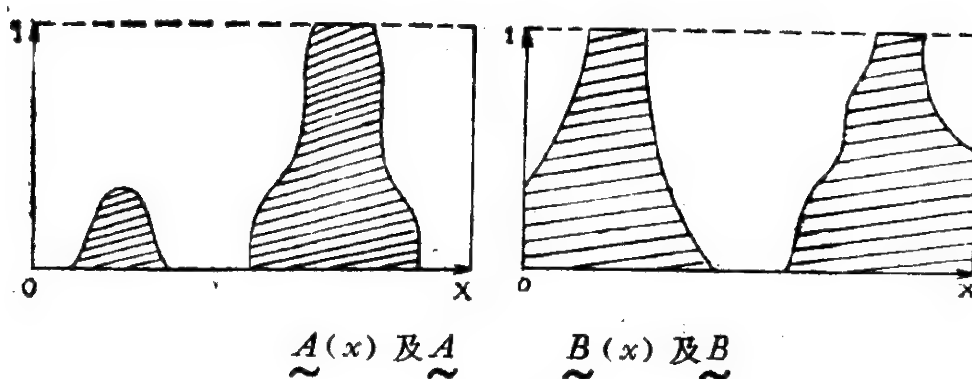
$$\underset{\sim}{A}^c(x)=1-\underset{\sim}{A}(x).$$

其几何表示见图 1.2.3~5。显然，可定义 F 集合的无限交与无限并的运算如下：

$$\left(\bigcup_{t\in T}\underset{\sim}{A}_t\right)(x)=\bigvee_{t\in T}\underset{\sim}{A}_t(x)=\sup_{t\in T}\underset{\sim}{A}_t(x),$$

$$\left(\bigcap_{t\in T}\underset{\sim}{A}_t\right)(x)=\bigwedge_{t\in T}\underset{\sim}{A}_t(x)=\inf_{t\in T}\underset{\sim}{A}_t(x),$$

其中 $\underset{\sim}{A}_t\in\mathscr{F}(X)(t\in T)$，显然并、交及补的结果皆属于 $\mathscr{F}(X)$。



$\underset{\sim}{A}(x)$ 及 $\underset{\sim}{A}$　　$\underset{\sim}{B}(x)$ 及 $\underset{\sim}{B}$

图 1.2.2

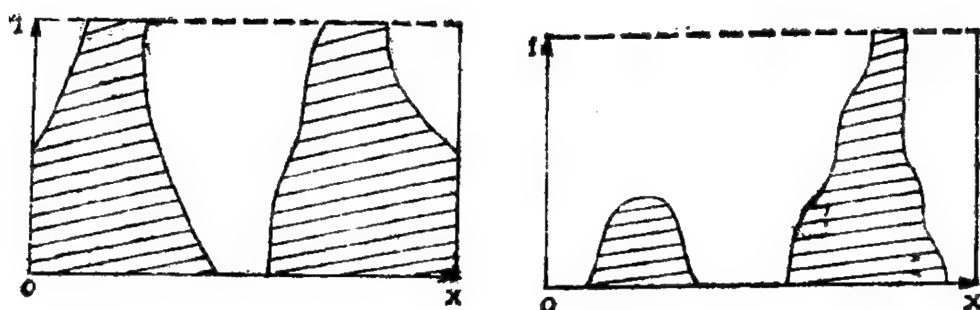


图 1.2.3　$(\underset{\sim}{A}\cup\underset{\sim}{B})(x)$ 及 $\underset{\sim}{A}\cup\underset{\sim}{B}$　图 1.2.4 $(\underset{\sim}{A}\cap\underset{\sim}{B})(x)$ 及 $\underset{\sim}{A}\cap\underset{\sim}{B}$

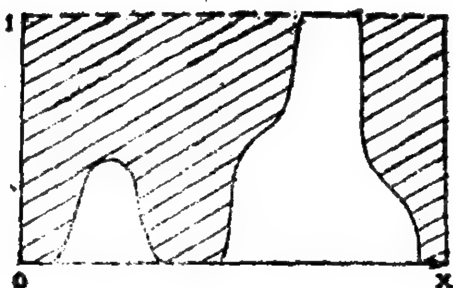


图 1.2.5　$\underset{\sim}{A}^c(x)$ 及 $\underset{\sim}{A}^c$

例 1.2.5　设 $X=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5\}$，而

$$\underset{\sim}{A}=0.4/\omega_1+0.2/\omega_2+1/\omega_4;$$

$$\underset{\sim}{B}=0.3/\omega_1.$$

则

$$\underset{\sim}{A}\cup\underset{\sim}{B}=0.4\vee0.3/\omega_1+0.2\vee0/\omega_2+1\vee0/\omega_4$$
$$=0.4/\omega_1+0.2/\omega_2+1/x_4;$$
$$\underset{\sim}{A}\cap\underset{\sim}{B}=0.4\wedge0.3/\omega_1+0.2\wedge0/\omega_2+1\wedge0/\omega_4$$
$$=0.3/\omega_1+0/\omega_2+0/\omega_4=0.3/\omega_1;$$
$$\underset{\sim}{A}^c=(1-0.4)/\omega_1+(1-0.2)/\omega_2+(1-0)/\omega_3$$
$$+(1-1)/\omega_4+(1-0)/\omega_5$$
$$=0.6/\omega_1+0.8/\omega_2+1/\omega_3+1/\omega_5.$$

例 1.2.6　设 $X=[0,1]$，$\underset{\sim}{A}(\omega)=\omega$.

则　$\underset{\sim}{A}^c(\omega)=1-\omega$;

$$(\underset{\sim}{A}\cup\underset{\sim}{A}^c)(\omega)=\begin{cases}1-\omega, & 若\ \omega\leqslant\frac{1}{2},\\ \omega, & 若\ \omega>\frac{1}{2};\end{cases}$$

$$(\underset{\sim}{A}\cap\underset{\sim}{A}^c)(\omega)=\begin{cases}\omega, & \text{若 } \omega\leqslant\frac{1}{2},\\ 1-\omega, & \text{若 } \omega>\frac{1}{2};\end{cases}$$

例 1.2.7 设 $X=[0,100]$，表示某年龄集，$\underset{\sim}{A}$ 和 $\underset{\sim}{B}$ 分别表示"年老"和"年轻"，其隶属函数如例 1.2.3。于是 $\underset{\sim}{A}\cup\underset{\sim}{B}$ 表示"年老或年轻"，$\underset{\sim}{A}\cap\underset{\sim}{B}$ 表示"又老又年轻"，$\underset{\sim}{A}^c$ 表示"不年轻"，其隶属函数分别为

$$(\underset{\sim}{A}\cup\underset{\sim}{B})(\omega)=\begin{cases}1 & 0\leqslant\omega<25\\ \left[1+\left(\frac{\omega-25}{5}\right)^2\right]^{-1}, & 25<\omega\leqslant 51\\ \left[1+\left(\frac{\omega-50}{5}\right)^{-2}\right]^{-1}, & 51<\omega\leqslant 100\end{cases}$$

$$(\underset{\sim}{A}\cap\underset{\sim}{B})(\omega)=\begin{cases}0 & 0\leqslant\omega\leqslant 50\\ \left[1+\left(\frac{\omega-50}{5}\right)^{-2}\right]^{-1} & 50<\omega\leqslant 51\\ \left[1+\left(\frac{\omega-25}{5}\right)^{2}\right]^{-1} & 51<\omega\leqslant 100\end{cases}$$

$$\underset{\sim}{A}^c(\omega)=\begin{cases}0, & 0\leqslant\omega\leqslant 25\\ 1-\left[1+\left(\frac{\omega-25}{5}\right)^2\right]^{-1}, & 25<\omega\leqslant 100\end{cases}$$

定理 1.2.1 $(\mathscr{F}(X),\cup,\cap;c,\phi,X)$具有以下性质：

(1) $\phi\subset\underset{\sim}{A}\subset X$ （最大、最小 F 集的存在性）

(2) $\underset{\sim}{A}\subset\underset{\sim}{A}$ （自反律）

(3) 若 $\underset{\sim}{A}\subset\underset{\sim}{B}$，$\underset{\sim}{B}\subset\underset{\sim}{C}$，则 $\underset{\sim}{A}\subset\underset{\sim}{C}$ （传递律）

(4) 若 $\underset{\sim}{A}\subset\underset{\sim}{B}$，$\underset{\sim}{B}\subset\underset{\sim}{A}$，则 $\underset{\sim}{A}=\underset{\sim}{B}$ （对称律）

(5) $\underset{\sim}{A}\cup\underset{\sim}{B}=\underset{\sim}{B}\cup\underset{\sim}{A}$，$\underset{\sim}{A}\cap\underset{\sim}{B}=\underset{\sim}{B}\cap\underset{\sim}{A}$ （交换律）

(6) $(\underset{\sim}{A}\cup\underset{\sim}{B})\cup\underset{\sim}{C}=\underset{\sim}{A}\cup(\underset{\sim}{B}\cup\underset{\sim}{C})$，

$(\underset{\sim}{A}\cap\underset{\sim}{B})\cap\underset{\sim}{C}=\underset{\sim}{A}\cap(\underset{\sim}{B}\cap\underset{\sim}{C})$； （结合律）

(7) $\underset{\sim}{A}\cup(\underset{\sim}{B}\cap\underset{\sim}{C})=(\underset{\sim}{A}\cup\underset{\sim}{B})\cap(\underset{\sim}{A}\cup\underset{\sim}{C})$，

$\underset{\sim}{A}\cap(\underset{\sim}{B}\cup\underset{\sim}{C})=(\underset{\sim}{A}\cap\underset{\sim}{B})\cup(\underset{\sim}{A}\cap\underset{\sim}{C})$， （分配律）

(8) $\underset{\sim}{A}\cup(\underset{\sim}{A}\cap\underset{\sim}{B})=\underset{\sim}{A}$，$\underset{\sim}{A}\cap(\underset{\sim}{A}\cup\underset{\sim}{B})=\underset{\sim}{A}$ （吸收律）

(9) $\underset{\sim}{A}\cup\underset{\sim}{A}=\underset{\sim}{A}$，$\underset{\sim}{A}\cap\underset{\sim}{A}=\underset{\sim}{A}$， （幂等律）

(10) $(\underset{\sim}{A}^c)^c=\underset{\sim}{A}$ （对合律）

(11) $X\cap\underset{\sim}{A}=\underset{\sim}{A}$，$X\cup\underset{\sim}{A}=X$

$\phi\cup\underset{\sim}{A}=\underset{\sim}{A}$，$\phi\cap\underset{\sim}{A}=\phi$， （两极律）

(12) $(\underset{\sim}{A}\cup\underset{\sim}{B})^c=\underset{\sim}{A}^c\cap\underset{\sim}{B}^c$，

$(\underset{\sim}{A}\cap\underset{\sim}{B})^c=\underset{\sim}{A}^c\cup\underset{\sim}{B}^c$， （对偶律）

若 $\underset{\sim}{A}_t\in\mathscr{F}(X)$，$(t\in T)$，则(7)和(12)有一般形式：

$$(7)'\quad \underset{\sim}{B}\cup\left(\bigcap_{t\in T}\underset{\sim}{A}_t\right)=\bigcap_{t\in T}(\underset{\sim}{B}\cup\underset{\sim}{A}_t),$$

$$\underset{\sim}{B}\cap\left(\bigcup_{t\in T}\underset{\sim}{A}_t\right)=\bigcup_{t\in T}(\underset{\sim}{B}\cap\underset{\sim}{A}_t);$$

$$(12)'\quad \left(\bigcup_{t\in T}\underset{\sim}{A}_t\right)^c=\bigcap_{t\in T}\underset{\sim}{A}_t^c,$$

$$\left(\bigcap_{t\in T}\underset{\sim}{A}_t\right)^c=\bigcup_{t\in T}\underset{\sim}{A}_t^c.$$

证 直接验证即可。我们来证明对偶律。

$$\begin{aligned}(\underset{\sim}{A}\cap\underset{\sim}{B})^c(x)&=1-(\underset{\sim}{A}\cap\underset{\sim}{B})(x)\\&=1-\min(\underset{\sim}{A}(x),\ B(x))\\&=\max(1-\underset{\sim}{A}(x),1-\underset{\sim}{B}(x))\\&=\max(\underset{\sim}{A}^c(x),\underset{\sim}{B}^c(x))\\&=(\underset{\sim}{A}^c\cup\underset{\sim}{B}^c)(x),\end{aligned}$$

则

$$(\underset{\sim}{A}\cap\underset{\sim}{B})^c=\underset{\sim}{A}^c\cup\underset{\sim}{B}^c.$$

其它的请读者自己验证。

熟悉经典集合的读者会发现这里没有“互补律”，即 $\underset{\sim}{A}\cup\underset{\sim}{A}^c=X$，$\underset{\sim}{A}\cap\underset{\sim}{A}^c=\phi$ 一般不再成立。例如例 1.3.2 中之 $\underset{\sim}{A}(\omega)=\omega$，其 $\underset{\sim}{A}\cup\underset{\sim}{A}^c\neq X$，$\underset{\sim}{A}\cap\underset{\sim}{A}^c\neq\phi$。特别

$$(\underset{\sim}{A}\cup\underset{\sim}{A}^c)\cdot\left(\frac{1}{2}\right)=(\underset{\sim}{A}\cap\underset{\sim}{A}^c)\left(\frac{1}{2}\right)=\frac{1}{2}.$$

从而可见 F集合不满足互补律。为了克服这一难点，考虑运算 $\cap$ 和 $\cup$ 在此为$\wedge$与$\vee$运算，我们将互补律改为

$$\underset{\sim}{A}\cup\underset{\sim}{A}^c=\max(\underset{\sim}{A},\underset{\sim}{A}^c),$$

$$\underset{\sim}{A}\cap\underset{\sim}{A}^c=\min(\underset{\sim}{A},\underset{\sim}{A}^c).$$

于是，互补律对于$\mathscr{F}(X)$也成立。本例中互补律显然成立。

例 1.2.8 设 $\underset{\sim}{A}(\omega)=0.2$，于是 $\underset{\sim}{A}^c(\omega)=0.8$则有

$$(\underset{\sim}{A}\cap\underset{\sim}{A}^c)(\omega)=0.2\wedge0.8=0.2$$

$$(\underset{\sim}{A}\cup\underset{\sim}{A}^c)(\omega)=0.2\vee0.8=0.8$$

即

$$\underset{\sim}{A}\cap\underset{\sim}{A}^c=\underset{\sim}{A},\quad \underset{\sim}{A}\cup\underset{\sim}{A}^c=\underset{\sim}{A}^c,$$

所以 $\underset{\sim}{A}$，$\underset{\sim}{A}^c$ 之间满足互补律。

为了方便，今后用记号$(\mathscr{F}(X,\vee,\wedge;[0,1])$表示由$X$构成的某些 F 集合的全体，其上定义了$\vee$，$\wedge$和逆运算，且隶属函数的值域为$[0,1]$。

定义 1.2.4 若$(\mathscr{F}(X),\wedge,\vee;[0,1])$满足

(1) $1\in\mathscr{F}(X)$；

(2) 若 $\underset{\sim}{A}_i\in\mathscr{F}(X)$，$i=1,2,\cdots$，则 $\bigvee_i\underset{\sim}{A}_i\in\mathscr{F}(X)$；

(3) 若 $\underset{\sim}{A}\in\mathscr{F}(X)$，则 $1-\underset{\sim}{A}\in\mathscr{F}(X)$。

则称$\mathscr{F}(X)$为σ软代数(Solf algebra)或伪σ代数。

例1.2.9 设论域$X=\{x_1,x_2\}$，隶属函数的取值集为$M=\left\{0,\frac{1}{2},1\right\}$。那么$X$的一切$F$子集的集系$\mathscr{F}(X)$为

$$\{[(x_1,0),(x_2,0)],[(x_1,0),(x_2,0.5)],$$
$$[(x_1,0),(x_2,1)],[(x_1,0.5),(x_2,0)],$$
$$[(x_1,0.5),(x_2,0.5)],[(x_1,0.5),(x_2,1)],$$
$$[(x_1,1),(x_2,0)],\ [(x_1,1),(x_2,0.5)],$$
$$[(x_1,1),(x_2,1)]\}.$$

$\mathscr{F}(X)$中共含有3^2个元素。

由于

(1) $[(x_1,1),(x_2,1)]\in\mathscr{F}(X)$，即$1\in\mathscr{F}(X)$；

(2) $[(x_1,0.5),(x_2,1)],[(x_1,1),(x_2,0)]\in\mathscr{F}(X)$

而

$$[(x_1,0.5),(x_2,1)]\vee[(x_1,1),(x_2,0)]$$
$$=[(x_1,0.5\vee 1),(x_2,1\vee 0)]$$
$$=[(x_1,1),(x_2,1)]\in\mathscr{F}(X),$$

等等，可得出$\underset{\sim}{A}$，$\underset{\sim}{B}\in\mathscr{F}(X)$，则$\underset{\sim}{A}\vee\underset{\sim}{B}\in\mathscr{F}(X)$。

(3) $[(x_1,0.5),(x_2,1)]\in\mathscr{F}(X)$，从而有

$$[(x_1,0.5),(x_2,1)]^c$$
$$=1-[(x_1,0.5),(x_2,1)]$$
$$=[(x_1,0.5),(x_2,0)]\in\mathscr{F}(X)$$

等等，得到，若$\underset{\sim}{A}\in\mathscr{F}(X)$则$1-\underset{\sim}{A}\in\mathscr{F}(X)$。

由(1)-(3)验证，可知$\mathscr{F}(x)$为一σ软代数。

和概率论中所考虑的情形一样，当X为一有限集时，$\mathscr{F}(X)$取其一切F子集为元素，但在X为无限集时，不一定

要取 X 的一切 F 子集组成 $\mathscr{F}(X)$，可取一部分组成集系 $\mathscr{X}$，但 $\mathscr{X}$ 必须是一 σ 软代数。

§3　F 集合的分解定理

为了研究 F 子集与经典集合之间的关系，我们讨论 F 集合的分解定理。为此先引入一些概念。

定义 1.3.1　设 $\underset{\sim}{A}\in F(X)$，记

$$(\underset{\sim}{A})_\alpha=\underset{\sim}{A}_\alpha=\{x\mid \underset{\sim}{A}(x)\geqslant\alpha\},\quad (0\leqslant\alpha\leqslant1)$$

称 A_α 为 F 子集 $\underset{\sim}{A}$ 的 α 水平集（截集），如图 1.3.1所示。称

$$A_0=\{x;\underset{\sim}{A}(x)>0\}=\text{supp}\,\underset{\sim}{A}$$

为 $\underset{\sim}{A}$ 的支集。称

$$A_1=\{x;\underset{\sim}{A}(x)=1\}$$

为 $\underset{\sim}{A}$ 的核。

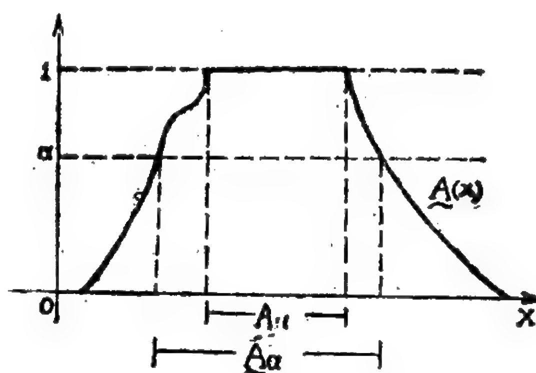


图 1.3.1　$\underset{\sim}{A}$ 的 α 水平集

例 1.3.1　我们考虑一个具体问题。设某学习班有 x_1，x_2，x_3，x_4，x_5，x_6 六位同学，在某次考试中成绩如下：

x_1	100 分	隶属度为 1
x_2	92 分	隶属度为 0.92
x_3	35 分	隶属度为 0.35
x_4	68 分	隶属度为 0.68
x_5	82 分	隶属度为 0.82
x_6	55 分	隶属度为 0.55

现在我们考虑"及格"(60分以上)"优秀"(90分以上)"良好"

(80分以上)都有哪些人？显然有

"及格"者集合 $A_{0.6}=\{x_1,x_2,x_4,x_5\}$，

"良好"者集合 $A_{0.8}=\{x_1,x_2,x_5\}$

"优秀"者集合 $A_{0.9}=\{x_1,x_2\}$

这三个集合就是按不同水平确定的普通集合，这些普通集合是对原来的 F 集 $\underset{\sim}{A}$ 的隶属度先确定一个水平 $\alpha(0\leqslant\alpha\leqslant1)$之后，再把隶属度 $\underset{\sim}{A}(x)\geqslant\alpha$ 的元素挑选出来而得到的。即就是由

$$\underset{\sim}{A}=1/x_1+0.92/x_2+0.35/x_3+0.68/x_4+0.82/x_5+0.55/x_6$$

分别就确定水平 $\alpha=0.6$，0.8，0.9而得到

$$A_{0.6}=\{x_1,x_2,x_4,x_5\}$$
$$A_{0.8}=\{x_1,x_2,x_5\}$$
$$A_{0.9}=\{x_1,x_2\}$$

显然

$$A_0=\{x_1,x_2,x_3,x_4,x_5,x_6\}=X,$$
$$A_1=\{x_1\}.$$

由此可见，F 集合 $\underset{\sim}{A}$ 的支集 $A_0=X$，且 $\alpha<\beta$时有$A_\beta\subset\underset{\sim}{A}_\alpha\subset X$。因此经典集族

$$\{A_\alpha;0<\alpha\leqslant1\}$$

表示一个边界不确定的集合。随着 α 从 1 下降趋于0(不到达0)，A_α 从 $\underset{\sim}{A}$ 的核 A_1 扩张到 $\underset{\sim}{A}$ 的支集 A_0。从而 F集合$\underset{\sim}{A}$是一个边界游移的集合。我们只能在某个 α 水平的意义下，认为 x 属于 $\underset{\sim}{A}$，还是不属于 $\underset{\sim}{A}$，而这个 α在此为隶属度。比如 $x\in A_\alpha$，我们称在 α 水平下 $x\in\underset{\sim}{A}$；而 $x\notin A_\alpha$，我们说在α 水平下 x 不属于 $\underset{\sim}{A}$。

定理 1.3.1 $\underset{\sim}{A}$ 的水平集具有以下性质：

(1) $(\underset{\sim}{A}\cap\underset{\sim}{B})_\alpha=A_\alpha\cap B_\alpha$;

(2) $(\underset{\sim}{A}\cup\underset{\sim}{B})_\alpha=A_\alpha\cup B_\alpha$;

(3) $\alpha,\beta\in[0,1]$，且 $\alpha<\beta$ 则 $A_\beta\subseteq A_\alpha$。

证
$$\begin{aligned}(\underset{\sim}{A}\cap\underset{\sim}{B})_\alpha&=\{x; (\underset{\sim}{A}\cap\underset{\sim}{B})(x)\geqslant\alpha\}\\&=\{x;\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x)\geqslant\alpha\}\\&=\{x; \underset{\sim}{A}(x)\geqslant\alpha\}\cap\{x;\underset{\sim}{B}(x)\geqslant\alpha\}\\&=A_\alpha\cap B_\alpha,\end{aligned}$$

$$\begin{aligned}(\underset{\sim}{A}\cup\underset{\sim}{B})_\alpha&=\{x;(\underset{\sim}{A}\cup\underset{\sim}{B})(x)\geqslant\alpha\}\\&=\{x;\underset{\sim}{A}(x)\vee\underset{\sim}{B}(x)\geqslant\alpha\}\\&=\{x;\underset{\sim}{A}(x)\geqslant\alpha\}\cup\{x;\underset{\sim}{B}(x)\geqslant\alpha\}\\&=A_\alpha\cup B_\alpha.\end{aligned}$$

至于 $A_\beta\subseteq A_\alpha$ 则显然。

定理 1.3.2 若$\{\underset{\sim}{A}_t; t\in T\}\subset\mathscr{F}(X)$，则

(1) $\left(\bigcup\limits_{t\in T}\underset{\sim}{A}_t\right)_\alpha\subset\bigcup\limits_{t\in T}(\underset{\sim}{A}_t)_\alpha$,

(2) $\left(\bigcap\limits_{t\in T}\underset{\sim}{A}_t\right)_\alpha=\bigcap\limits_{t\in T}(\underset{\sim}{A}_t)_\alpha$.

证 若 $x\in\bigcup\limits_{t\in T}(\underset{\sim}{A}_t)_\alpha$，则存在 $t_0\in T$，使 $x\in(\underset{\sim}{A}_{t_0})_\alpha$，于是 $\underset{\sim}{A}_{t_0}(x)\geqslant\alpha$，即得 $\sup\limits_{t\in T}\underset{\sim}{A}_t(x)\geqslant\alpha$，故 $x\in\left(\bigcup\limits_{t\in T}\underset{\sim}{A}_t\right)_\alpha$，证得(1)。

必须指出(1)中“$\subset$”不能换为等式。

至于(2)式的证明，可类似于(1)推证得到。

例 1.3.2 令

$$\underset{\sim}{A}_n(x)=\frac{1}{2}\left(1-\frac{1}{n}\right),$$

则

$$\left(\bigcup_{n=1}^{\infty} \underset{\sim}{A}_n\right)(x)=\frac{1}{2},$$

于是

$$\left(\bigcup_{n=1}^{\infty} \underset{\sim}{A}_n\right)_{0.5}=X.$$

但是

$$(\underset{\sim}{A}_n)_{0.5}=\phi \quad (n\geqslant 1)$$

因此

$$\bigcup_{n=1}^{\infty}(\underset{\sim}{A}_n)_{0.5}=\phi,$$

从而

$$\left(\bigcup_{n=1}^{\infty} \underset{\sim}{A}_n\right)_{0.5}\neq \bigcup_{n=1}^{\infty}(\underset{\sim}{A}_n)_{0.5}.$$

定理 1.3.3 设 $\underset{\sim}{A}\in\mathscr{F}(X)$，$\{\alpha_t;t\in T\}\subset[0,1]$，则

$$A_\alpha=\bigcap_{t\in T} A_{\alpha_t}$$

$$A_\beta\supset\bigcup_{t\in T} A_{\alpha_t},$$

其中 $\alpha=\bigvee_{t\in T}\alpha_t$，$\beta=\bigwedge_{t\in T}\beta_t$.

证 由于

$$\begin{aligned}A_\alpha&=\left\{x;\underset{\sim}{A}(x)\geqslant\bigvee_{t\in T}\alpha_t\right\}\\&=\bigcap_{t\in T}\{x;\underset{\sim}{A}(x)\geqslant\alpha_t\}=\bigcap_{t\in T}A_{\alpha_t}\end{aligned}$$

又

$$\begin{aligned}A_\beta&=\left\{x;\underset{\sim}{A}(x)\geqslant\bigwedge_{t\in T}\alpha_t\right\}\\&\supset\bigcup_{t\in T}\{x;\underset{\sim}{A}(x)\geqslant\alpha_t\}=\bigcup_{t\in T}A_{\alpha_t},\end{aligned}$$

证得定理结论。

定理 1.3.4 对于任意 $\underset{\sim}{A} \in \mathscr{F}(X)$，有

$$A_\alpha = \bigcap_{\lambda<\alpha} A_\lambda$$

证 由于

$$\begin{aligned} A_\alpha &= \{x : \underset{\sim}{A}(x) \geqslant \alpha\} \\ &= \bigcap_{\lambda<\alpha} \{x : \underset{\sim}{A}(x) \geqslant \alpha\} = \bigcap_{\lambda<\alpha} A_\lambda \end{aligned}$$

得证。

定义 1.3.2 设 $\alpha \in [0,1]$，$\underset{\sim}{A} \in \mathscr{F}(X)$，则称

$$(\alpha \underset{\sim}{A})(x) = \alpha \wedge \underset{\sim}{A}(x)$$

为 α 与 $\underset{\sim}{A}$ 的数积。

α 与 $\underset{\sim}{A}$ 的数积是一个 F 子集。

定理 1.3.5 （F 集合的分解定理）对于任意 $\underset{\sim}{A} \in \mathscr{F}(X)$，有

$$\underset{\sim}{A} = \bigcup_{\alpha \in [0,1]} \alpha A_\alpha,$$

若 R_0 为[0,1]中的有理点集，则

$$\underset{\sim}{A} = \bigcup_{\alpha \in R_0} \alpha A_\alpha.$$

证 因为

$$A_\alpha(x) = \begin{cases} 1, & x \in A_\alpha, \\ 0, & x \overline{\in} A_\alpha. \end{cases}$$

则

$$\begin{aligned} \Big(\bigcup_{\alpha \in [0,1]} \alpha A_n\Big)(x) &= \sup_{0 \leqslant \alpha \leqslant 1} \alpha \cdot A_\alpha(x) \\ &= \sup_{x \in A_\alpha} \alpha = \sup_{\alpha \leqslant \underset{\sim}{A}(x)} \alpha = \underset{\sim}{A}(x). \end{aligned}$$

其它形式类似可证。

我们用图来说明这个定理。在图 1.3.1 中画出三个不同水平α，α_1，α_2 的隶属函数$\alpha A_\alpha(x)$，$\alpha_1 A_{\alpha_1}(x)$，$\alpha_2 A_{\alpha_2}(x)$。

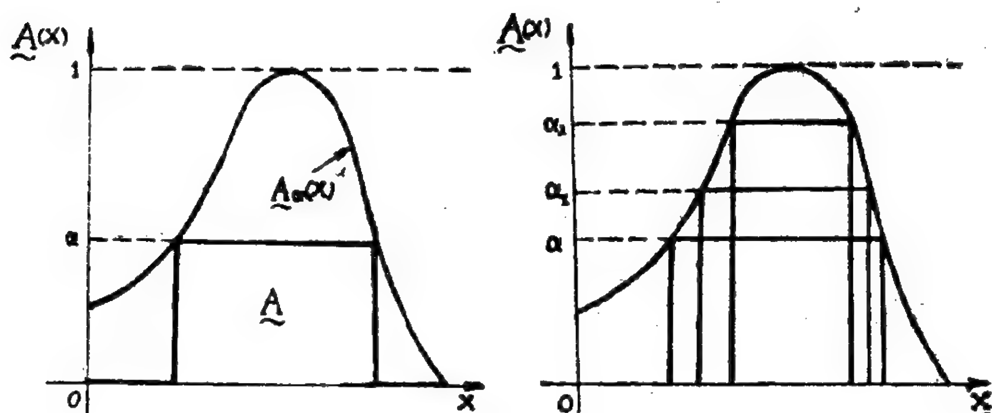


图 1.3.2

设想当α取遍[0,1]所有值时，$\cup \alpha A_\alpha$ 按 F 集求并运算法则，也就是取各$\alpha \in [0,1]$点隶属函数的最大值，再连成一条曲线，这自然就与$\underset{\sim}{A}(x)$的曲线重合，这就是分解定理。

§4 扩张原理

假设给定了两个集合 X 和 Y，且给定一映射

$$\begin{cases} f: & X \to Y, \\ & x \to y = f(x), x \in X, \quad y \in Y. \end{cases}$$

如果在 X 上给定一经典子集 A，则可以通过映射 f 得到一个集合 $B = f(A)$，且 $B \subseteq Y$。但是若在 X 给定的集为一个 F 集 $\underset{\sim}{A}$，那么经过 f 映射之后变成什么呢？为了解答此问题，Zadeh 在 1975 年引入"扩张原理"，它将经典映射推广至 F 集之间。

众所周知，当我们知道 X 和 Y的幂集$\mathscr{P}(X)$及$\mathscr{P}(Y)$时，

f 能在幂集之间进行：

$$f: \mathscr{P}(X) \to \mathscr{P}(Y),$$

且

$$f(A)=\{y;\ y=f(x),\ x\in A\},\ A\in\mathscr{P}(X).$$

$$f^{-1}(B)=\{x\mid f(x)=y,\ y\in B\},\ B\in\mathscr{P}(Y).$$

对于模糊集，我们有如下定义：

定义 1.4.1（扩张原理） 设 $f: X\to Y$，记

$$f(\underset{\sim}{A})(y)=\bigvee\{\underset{\sim}{A}(x);\ x\in f^{-1}(y)\},$$

称 $f: \mathscr{F}(X)\to\mathscr{F}(Y)$ 为直接映射。

扩张原理可作如下解释：$\underset{\sim}{A}$ 经映射 f 后变成象 $f(\underset{\sim}{A})$ 时，其隶属函数可无保留的传递过去，亦即经过映射后，模糊子集 $\underset{\sim}{A}$ 和 $f(\underset{\sim}{A})$ 论域中相应元素的隶属函数保持不变。若不是单值映射时，则规定象的隶属度取最大值。如图 1.4.1 及 1.4.2 所示。图 1.4.1 表示单值映射后的情形。$\underset{\sim}{A}$ 经过映射后

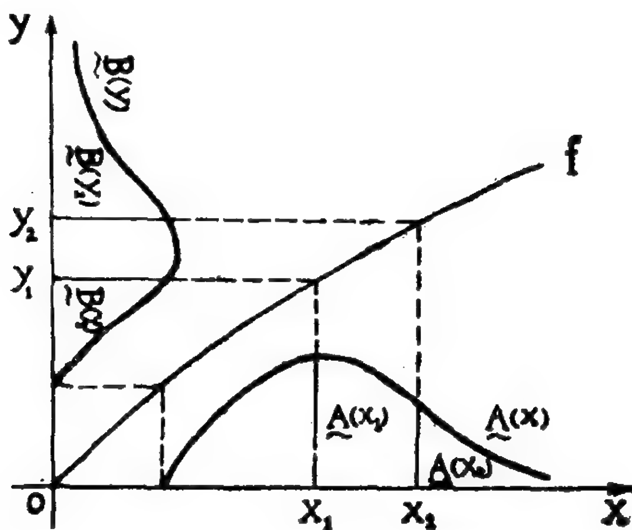


图 1.4.1 单值映射后隶属度保持不变

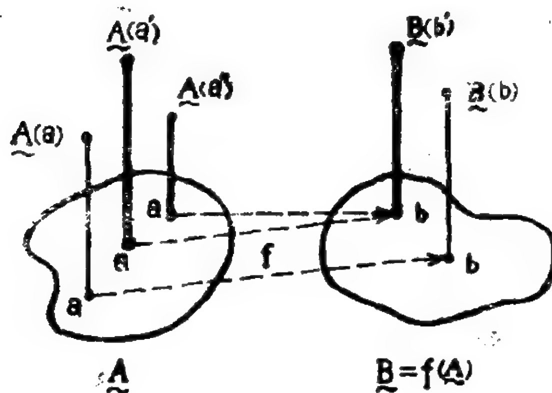


图 1.4.2 非单值映射后隶属度取最大者

成为 $\underset{\sim}{B}=f(\underset{\sim}{A})$，其隶属函数$\underset{\sim}{A}(\cdot)$变成为$\underset{\sim}{B}(\cdot)$，但各相应点的隶属度不变，如

x_1点之相应点 y_1，$\underset{\sim}{B}(y_1)=\underset{\sim}{A}(x_1)$，

x_2点之相应点 y_2，$\underset{\sim}{B}(y_2)=\underset{\sim}{A}(x_2)$

等等。若映射为非单值的如图 1.4.2 所示，设

$$\underset{\sim}{A}=\underset{\sim}{A}(a)/a+\underset{\sim}{A}(a')/a'+\underset{\sim}{A}(a'')/a''$$

经 f 映射后

$$f(\underset{\sim}{A})=\underset{\sim}{B}=\underset{\sim}{B}(b)/b+\underset{\sim}{B}(b')/b'$$

其中　a'，a'' 经 f 变到 b'，a 对应 b，

$$\underset{\sim}{B}(b)=\underset{\sim}{A}(a),$$

$$\underset{\sim}{B}(b')=\underset{\sim}{A}(a')\vee\underset{\sim}{A}(a'').$$

定理 1.4.1　直接映射 f 满足

(1) $f(\phi)=\phi$,

(2) $f\left(\bigcup\limits_{t\in T}\underset{\sim}{A}_t\right)=\left(\bigcup\limits_{t\in T}f(\underset{\sim}{A}_t)\right)$

(3) $f^{-1}(\underset{\sim}{B}^c)=(f^{-1}(\underset{\sim}{B}))^c$.　　　　$\underset{\sim}{B}\in\mathscr{F}(Y)$

（称 f 为序同态的）。而且

$$f^{-1}(\underset{\sim}{B})(x)=\underset{\sim}{B}(f(x))=(\underset{\sim}{B}\circ f)(x).$$

证 显然 $f(\phi)=\phi$。其次设 $\underset{\sim}{A}_t\in\mathscr{F}(X)$，$(t\in T)$，则

$$\begin{aligned}
f\Big(\bigcup_{t\in T}\underset{\sim}{A}_t\Big)(y)&=\bigvee\Big\{\bigvee_{t\in T}\underset{\sim}{A}_t(x);\ x\in f^{-1}(y)\Big\}\\
&=\bigvee_{t\in T}\bigvee\{\underset{\sim}{A}_t(x);\ x\in f^{-1}(y)\}\\
&=\bigvee_{t\in T}f(\underset{\sim}{A}_t)(y)\\
&=\Big(\bigcup_{t\in T}f(\underset{\sim}{A}_t)\Big)(y).
\end{aligned}$$

由于

$$\begin{aligned}
f^{-1}(\underset{\sim}{B})&=\cap\{\underset{\sim}{A};f(\underset{\sim}{A}^c)\subset\underset{\sim}{B}^c\}\\
&=\cap[\underset{\sim}{A};\bigvee\{\underset{\sim}{A}^c(x);x\in f^{-1}(y)\}\\
&\quad\leqslant\underset{\sim}{B}^c(y),(y\in Y)]\\
&=\cap[\underset{\sim}{A};\bigwedge\{\underset{\sim}{A}(x);x\in f^{-1}(y)\}\\
&\quad\geqslant\underset{\sim}{B}(y),(y\in Y)]\\
&=\underset{\sim}{B}(f(x))
\end{aligned}$$

则 $f^{-1}(\underset{\sim}{B}^c)=(f^{-1}(\underset{\sim}{B}))^c$ 得证。

其中“$\circ$”表示复合之意。设

$$f:\ X\to Y,\quad g:\ Y\to Z$$

为映射，f 和 g 的复合映射定义为

$$g\circ f: X\to Z.$$
$$(g\circ f)(x)=g(f(x)).$$

从而不难明白 $\underset{\sim}{B}\circ f$ 之含义。

定理 1.4.2 直接映射 f 和其逆映射 f^{-1} 之间有

(1) $f(f^{-1}(\underset{\sim}{B}))\subset\underset{\sim}{B}$;

(2) $f^{-1}(f(\underset{\sim}{A}))\supset\underset{\sim}{A}$,

其中 $\underset{\sim}{A}\in\mathscr{F}(X)$, $\underset{\sim}{B}\in\mathscr{F}(Y)$.

证 由于

$$f(f^{-1}(\underset{\sim}{B}))(y)=\bigvee\{f^{-1}(\underset{\sim}{B})(x);\ x\in f^{-1}(y)\}$$
$$=\bigvee\{\underset{\sim}{B}(f(x));\ x\in f^{-1}(y)\}\leqslant B(y)$$

则证(1)。又因

$$f^{-1}(f(\underset{\sim}{A}))(x)=f(\underset{\sim}{A})(f(x))$$
$$=\bigvee\{\underset{\sim}{A}(x);\ x\in f^{-1}(f(x))\}\geqslant\underset{\sim}{A}(x)$$

则证(2)。

例 1.4.1 设"小的"数的 F 集为 $\underset{\sim}{A}$:

$$\underset{\sim}{A}=\frac{1}{0}+\frac{1}{1}+\frac{0.8}{2}+\frac{0.6}{3}+\frac{0.4}{4}+\frac{0.2}{5}$$

而令 $f(u)=u^2$, 则 f(小的)为

$$f(\text{小的})=f(\underset{\sim}{A})=\underset{\sim}{A}^2$$
$$=\frac{1}{0}+\frac{1}{1}+\frac{0.8}{4}+\frac{0.6}{9}+\frac{0.4}{16}+\frac{0.2}{25}.$$

§5 F 数及其扩张运算

作为扩张原理的应用, 首先考虑的是不分明数 (以下简称为 F 数)。设以 R 表示全体实数, $\mathscr{F}(R)$ 表示实数域上的一切 F 集合之集.

定义 1.5.1 $\underset{\sim}{A}\in\mathscr{F}(R)$称为 F 数, 如果

(1) 存在 $x_0\in R$ 使 $\underset{\sim}{A}(x_0)=1$;

(2) $\forall\alpha\in(0,1]$, A_α 是闭区间.

定义 1.5.2 $\underset{\sim}{A}\in\mathscr{F}(R)$称为 F 凸的, 如果 $\forall x,y\in R$有

$$\underset{\sim}{A}(\lambda x+(1-\lambda)y)\geqslant\underset{\sim}{A}(x)\wedge\underset{\sim}{A}(y),\ (\lambda\in[0,1])$$

定理 1.5.1 设 $\underset{\sim}{A}$ 是 F 数，则

(1) $\underset{\sim}{A}$ 是凸的；

(2) 若 $\underset{\sim}{A}(x_0)=1$，则 $x\leqslant x_0$ 时，$\underset{\sim}{A}(x)$不减；$x\geqslant x_0$时，$\underset{\sim}{A}(x)$不增。

证 由于 $A_\alpha(\alpha\in[0,1])$是闭区间，$A_0=R$，即 $\forall\alpha\in[0,1]$，A_α是凸集，从而 $\underset{\sim}{A}$ 是凸的。

现设 $x_1<x_2\leqslant x_0$，令 $\alpha=\underset{\sim}{A}(x_1)$。由于 $\underset{\sim}{A}(x_0)=1$，则 $[x_1,x_0]\subset A_\alpha$。于是 $x_2\in A_\alpha$，从而 $\underset{\sim}{A}(x_2)\geqslant\alpha$，即 $\underset{\sim}{A}(x_1)\leqslant\underset{\sim}{A}(x_2)$。同理，如果 $x_0\leqslant x_1<x_2$，则可证 $\underset{\sim}{A}(x_2)\leqslant\underset{\sim}{A}(x_1)$。定理得证。

定义 1.5.3 设$*$为 R 上的二元运算. 扩张运算为

$$(\underset{\sim}{A}*\underset{\sim}{B})(z)=\bigvee_{z=x*y}(\underset{\sim}{A}(x)\wedge\underset{\sim}{A}(y))$$

其中$\underset{\sim}{A}$，$\underset{\sim}{B}\in\mathscr{F}(R)$。特别$*$为$+$，$-$，$\times$或$\div$时，称

$$(\underset{\sim}{A}+\underset{\sim}{B})(z)=\bigvee_{z=x+y}(\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(y))$$

$$(\underset{\sim}{A}-\underset{\sim}{B})(z)=\bigvee_{z=x-y}(\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(y))$$

$$(\underset{\sim}{A}\times\underset{\sim}{B})(z)=\bigvee_{z=x\cdot y}(\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(y))$$

$$(\underset{\sim}{A}\div\underset{\sim}{B})(z)=\bigvee_{z=x/y}(\underset{\sim}{A}(x\wedge)\underset{\sim}{B}(y))$$

为扩张加法、扩张减法、扩张乘法、扩张除法。

例 1.5.1 设 $X=Y=\{1,2,3\cdots,10\}$，$\underset{\sim}{2}$ 和 $\underset{\sim}{6}$ 分别表示“约为 2”和“约为 6”的 F 数，设

$$\underset{\sim}{2}=\frac{0.6}{1}+\frac{1}{2}+\frac{0.8}{3},$$

$$\underset{\sim}{6}=\frac{0.8}{5}+\frac{1}{6}+\frac{0.7}{7}.$$

则

$$\underset{\sim}{2}+\underset{\sim}{6}=\left(\frac{0.6}{1}+\frac{1}{2}+\frac{0.8}{3}\right)+\left(\frac{0.8}{5}+\frac{1}{6}+\frac{0.7}{7}\right)$$

$$=\bigvee_{z}\left(\frac{0.6\wedge 0.8}{1+5}+\frac{0.6\wedge 1}{1+6}+\frac{0.6\wedge 0.7}{1+7}\right.$$

$$+\frac{0.8\wedge 1}{2+5}+\frac{1\wedge 1}{2+6}+\frac{1\wedge 0.7}{2+7}+\frac{0.8\wedge 0.8}{3+5}$$

$$\left.+\frac{0.8\wedge 1}{3+6}+\frac{0.8\wedge 0.7}{3+7}\right)$$

$$=\bigvee_{z}\left[\frac{0.6}{6}+\frac{0.6}{7}+\frac{0.6}{8}+\frac{0.8}{7}+\frac{1}{8}+\frac{0.7}{9}\right.$$

$$\left.+\frac{0.8}{8}+\frac{0.8}{9}+\frac{0.7}{10}\right]$$

$$=\frac{0.6}{6}+\frac{0.6\vee 0.8}{7}+\frac{0.6\vee 1\vee 0.8}{8}$$

$$+\frac{0.7\vee 0.8}{9}+\frac{0.7}{10}$$

$$=\frac{0.6}{6}+\frac{0.8}{7}+\frac{1}{8}+\frac{0.8}{9}+\frac{0.7}{10}=\underset{\sim}{8}.$$

$$\underset{\sim}{6}-\underset{\sim}{2}=\bigvee_{z}\left[\frac{0.8\wedge 0.6}{5-1}+\frac{0.8\wedge 1}{5-2}+\frac{0.8\wedge 0.8}{5-3}\right.$$

$$+\frac{1\wedge 0.6}{6-1}+\frac{1\wedge 1}{6-2}+\frac{1\wedge 0.8}{6-3}+\frac{0.7\wedge 0.6}{7-1}$$

$$\left.+\frac{0.7\wedge 1}{7-2}+\frac{0.7\wedge 0.8}{7-3}\right]$$

$$=\bigvee_{z}\left[\frac{0.6}{4}+\frac{0.8}{3}+\frac{0.8}{2}+\frac{0.6}{5}+\frac{1}{4}\right.$$

$$+\frac{0.8}{3}+\frac{0.6}{6}+\frac{0.7}{5}+\frac{0.7}{4}\Big]$$

$$=\frac{0.8}{2}+\frac{0.8\vee 0.8}{3}+\frac{0.6\vee 0.7\vee 1}{4}$$

$$+\frac{0.6\vee 0.7}{5}+\frac{0.6}{6}$$

$$=\frac{0.8}{2}+\frac{0.8}{3}+\frac{1}{4}+\frac{0.7}{5}+\frac{0.6}{6}=\underset{\sim}{4}.$$

$$\underset{\sim}{2}\times\underset{\sim}{6}=\frac{0.6}{5}+\frac{0.6}{6}+\frac{0.6}{7}+\frac{0.8}{10}+\frac{1}{12}+\frac{0.7}{14}$$

$$+\frac{0.8}{15}+\frac{0.8}{18}+\frac{0.7}{21}=\underset{\sim}{12}.$$

$$\underset{\sim}{6}\div\underset{\sim}{2}=\frac{0.6}{5}+\frac{0.8}{2.5}+\frac{0.8}{5/3}+\frac{0.6}{6}+\frac{1}{3}+\frac{0.8}{2}$$

$$+\frac{0.6}{7}+\frac{0.7}{3.5}+\frac{0.7}{7/3}=\underset{\sim}{3}.$$

例 1.5.2 设

$$\underset{\sim}{2}=\int_1^2(x-1)/x+\int_2^3(3-x)/x,$$

则

$$\underset{\sim}{2}+\underset{\sim}{2}=\int_1^4\left(\frac{x}{2}-1\right)\Big/x+\int_4^6\left(3-\frac{x}{2}\right)\Big/x;$$

$$\underset{\sim}{2}-\underset{\sim}{2}=\int_{-2}^0\left(\frac{x}{2}+1\right)\Big/x+\int_0^2\left(1-\frac{x}{2}\right)\Big/x;$$

$$\underset{\sim}{2}\times\underset{\sim}{2}=\int_1^4(\sqrt{x}-1)/x+\int_4^9(3-\sqrt{x})/x.$$

$$\underset{\sim}{2}\div\underset{\sim}{2}=\int_{1/3}^1\left(3-\frac{4}{x+1}\right)\Big/x+\int_1^3\left(\frac{4}{x+1}-1\right)\Big/x.$$

不难验证，F数之间运算满足加法交换律、加法结合律、

乘法交换律、乘法结合律，加法存在 $\underset{\sim}{0}$ 元，乘法存在单位元 $\underset{\sim}{1}$．注意关于次分配律 $\underset{\sim}{A}\times(\underset{\sim}{J}+\underset{\sim}{K})\subseteq\underset{\sim}{A}\times\underset{\sim}{J}+\underset{\sim}{A}\times\underset{\sim}{K}$ 成立。而且 $\underset{\sim}{0}$ 不能作除数。顺便提及，定义 1.5.3 可由如下结论而得：设 $\underset{\sim}{A}$，$\underset{\sim}{B}$ 为两个 F 数，隶属函数分别为 $\underset{\sim}{A}(x)$，$\underset{\sim}{B}(x)$，则对于任意 $\alpha\in[0,1]$，任意 $x,y,z\in R$，有

$$\{z\,|\,z=x*y,\ x\in\underset{\sim}{A}_\alpha,y\in\underset{\sim}{B}_\alpha\}$$

$$=\left\{z\,\middle|\,\bigvee_{z=x*y}(\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x))\geqslant\alpha\right\}.$$

可见，F 数的运算关键在于求平集．这些内容我们将在第五章中用到。

第二章

不分明概率论的初等理论

不分明数学理论与概率论之间既有区别又有联系，这种联系产生了不分明概率论。特别是不分明集合理论与概率论的结合产生了不分明事件的概率问题，将是本章的主要论题.

不分明事件的概率问题首先由Zadeh作为隶属函数的数学期望给出，其后Klement提出公理化定义，Smets对它又作了一些修改。最近又有不少学者提出修正方案。本章将由浅入深的对这些加以介绍。

§1 引　言

不分明概率论应和经典概率论一样，是研究随机现象统计规律性的数学学科.显然，这里考虑的随机事件是具有不分明性质的随机事件。如明天天气很热的概率是多少？射击不多几次就命中目标的概率为何？所抽取的产品中次品大约为4个的概率是多大？老年人患慢性气管炎的概率是多少？诸如此类的问题可在人们的日常生活中遇到，它们要求的是某个事件的概率，可是这个事件“很热”、“不多几次”、“大约为4”、“老年人”等不是经典概率论中所讨论的事件，也不能用经典集合来刻划。通过第一章的学习，我们知道像这样的不分明概念所反映的事件可用F集合来表现. 这种事件我们在以后称之为F事件。因此，上面的诸多问题所涉及的是求F事件的概率问题。

例 2.1.1 设某人为办旅游事业，对某地一个特殊的夏天，询问一些旅游人的意见，意见分为三类"很暖"、"暖"及"不表态"，询问结果见表一。

表1

答案	答复的分数
很暖	0.4
暖	0.5
不表态	0.1

显然，每一份询问意见表是一次试验，而一个答案是一个试验的结果. 因此样本空间 Ω是明确的，而事件集合为

$$\mathscr{F}(\Omega)=\{\text{很暖，暖，不表态}\}.$$

其中每一个事件大约是由于概念的朦胧性带来了不分明性。如"很暖"和"暖"二者之间就无明确的界限，往往由被询问人的感觉认识来决定. 但是每一个事件又具有随机性，这些随机性是由那些变化的无意识性产生的。如询问结果中有十分之四的人回答"很暖"，那么"很暖"这一事件的概率为0.4。

由此可见，例中事件是不分明的，而概率是明确的。

例 2.1.2 掷一个骰子，考查出现较大点数的概率。

在掷一次骰子后到底出现什么样的点数这是事先无法断言的，这产生了随机性. 而出现什么样的点数算较大点数，这是不明确的，从而带来了不分明性. 掷一颗骰子出现什么点数是随机试验，每次试验一定会出现一个点数，因此样本空间 Ω 是明确的：

$$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$$

其中 ω_i 表示 i 点，$i=1,2,\cdots 6$。所关心的"出现较大点数"的事件，记为 $\underset{\sim}{A}$，是 Ω 的一个子集合. 但是什么样的样本点属于 $\underset{\sim}{A}$ 呢？我们可依照第一章 F 集合的理论，设

$$\underset{\sim}{A}=0.5/\omega_4+0.8/\omega_5+1/\omega_6.$$

其中分子表示隶属度，如 $0.8/\omega_5$ 表示“五点”隶属于“较大点数”的程度为 0.8。

有了以上分析，可见我们关心的是 $P(\underset{\sim}{A})$ 为何。其中 $\underset{\sim}{A}$ 是给定的，即对于 Ω 的元素 ω 其隶属于 $\underset{\sim}{A}$ 的程度 $\underset{\sim}{A}(x)$是已知的。

§2 F 事件

在经典概率论中，随机事件是用集合来刻划的。但在 F 概率论中，一个随机事件显然用 F 集合来表现. 随机试验的概念和经典概率论的概念一样，它应该在相同条件下能重复地进行；每次试验的可能结果不止一个，并且能事先明确试验的所有可能结果；最后是在进行一次试验之前不能确定哪一个结果会出现. 试验的一切可能结果 ω 之集，称为样本空间，记为 $\Omega=(\omega)$。

一试验中，除了这些最简单的事件(元素 ω 的单个事件)以外还有其它的随机事件，例如在 §1 例 2.1.2 中“较大点数”也是一个随机事件，它是由“出现 4 点”、“出现 5 点、“出现 6 点”这三个最简单的事件所组成，并且在“出现 4 点”时“较大点数 $\underset{\sim}{A}$”发生了，但却只有 0.5 的程度隶属于 $\underset{\sim}{A}$。同样，“出现 5 点”时 $\underset{\sim}{A}$发生了，但隶属的程度为 0.8；“出现 6 点”时隶属于 $\underset{\sim}{A}$ 的程度为 1. 因此，$\underset{\sim}{A}$ 所含的最简单事件之出现确定了 $\underset{\sim}{A}$的发生，但伴随一个隶属程度$\underset{\sim}{A}(\omega)$.这显然是一个 F 集合，我们称这样具有随机性的 F 集合为一个不分明事件，简称 F 事件。

F 事件是一个 F 集合，因此 F 事件间的运算关系亦如 F 集合的运算关系一样，不再一一叙述了(见第一章)。

设样本空间为Ω，其元素为ω，$\Omega=(\omega)$。Ω上一切F集合的全体记为$\mathscr{F}(\Omega)$。并设$\underset{\sim}{A},\underset{\sim}{B}\in\mathscr{F}(\Omega)$。

定义 2.2.1 设F事件$\underset{\sim}{A}$、$\underset{\sim}{B}\in\mathscr{F}(\Omega)$，若对于任意的$\omega\in\Omega$，有$\underset{\sim}{A}(\omega)\leqslant\underset{\sim}{B}(\omega)$，则称$F$事件$\underset{\sim}{A}$含于$\underset{\sim}{B}$，或说$\underset{\sim}{B}$包含$\underset{\sim}{A}$，并记作$\underset{\sim}{A}\subset\underset{\sim}{B}$。若对于任意的$\omega\in\Omega$有$\underset{\sim}{A}(\omega)=\underset{\sim}{B}(\omega)$，称两个$F$事年$\underset{\sim}{A}$与$\underset{\sim}{B}$相等，并记作$\underset{\sim}{A}=\underset{\sim}{B}$。

ϕ表示隶属函数恒为0的事件，称ϕ为不可能事件，用Ω表示隶属函数恒为1的事件，称Ω为必然事件。ϕ和Ω是F事件的极端情形。

定义 2.2.2 两个F事件$\underset{\sim}{A}$、$\underset{\sim}{B}$的并、交及逆事件$\underset{\sim}{A}\cup\underset{\sim}{B}$、$\underset{\sim}{A}\cap\underset{\sim}{B}$及$\underset{\sim}{A}^c$分别由隶属函数

$$(\underset{\sim}{A}\cup\underset{\sim}{B})(\omega)=\underset{\sim}{A}(\omega)\vee\underset{\sim}{B}(\omega)=\max(\underset{\sim}{A}(\omega),\underset{\sim}{B}(\omega)),$$

$$(\underset{\sim}{A}\cap\underset{\sim}{B})(\omega)=\underset{\sim}{A}(\omega)\wedge\underset{\sim}{B}(\omega)=\min(\underset{\sim}{A}(\omega),\underset{\sim}{B}(\omega)),$$

$$\underset{\sim}{A}^c(\omega)=1-\underset{\sim}{A}(\omega)$$

确定。

由此可见，F事件的理论实际上就是F集合的理论，为了给出F事件的概率我们必需给F事件下一新定义（见§3）。但目前有关F集合理论的内容足可以用来研究F事件。由第一章可知由Ω产生的$\mathscr{F}(\Omega)$形成一个σ软代数。F事件的运算为逐点对隶属函数作运算。从而F事件与F集合间形成一一对应关系，列举如下：

F概率论	F集合论
样本空间$\Omega=(\omega)$	论域$X=(x)$
基本事件ω	X的元素x
F随机事件$\underset{\sim}{A}$	F集合$\underset{\sim}{A}$
不可能事件ϕ	0

必然事件 Ω	1
$\underset{\sim}{A}$与 $\underset{\sim}{B}$ 至少发生一个：$\underset{\sim}{A}\cup\underset{\sim}{B}$	$\underset{\sim}{A}\vee\underset{\sim}{B}$
$\underset{\sim}{A}$与 $\underset{\sim}{B}$ 同时发生：$\underset{\sim}{A}\cap\underset{\sim}{B}$	$\underset{\sim}{A}\wedge\underset{\sim}{B}$
$\underset{\sim}{A}$ 的逆事件 $\underset{\sim}{A}^c$	$\underset{\sim}{A}^c$
$\underset{\sim}{A}$ 发生	$\underset{\sim}{A}(x)>0$

由于以上原因，在许多场合用 F 集合的语言叙述显得简练而自然。

定义 2.2.3 两个 F 事件 $\underset{\sim}{A}$、$\underset{\sim}{B}$ 称为互斥的，如果对任意 $\omega\in\Omega$ 有

$$\underset{\sim}{A}(\omega)\wedge\underset{\sim}{B}(\omega)=0.$$

这个定义说明 $\underset{\sim}{A}$、$\underset{\sim}{B}$ 互斥，则$\underset{\sim}{A}$、$\underset{\sim}{B}$ 的隶属函数 $\underset{\sim}{A}(\omega)$、$\underset{\sim}{B}(\omega)$在每一点 ω 不可能同时不为 0。

例 2.2.2 设 $\Omega=\{a,b,c,d,e\}$，而设

$$\underset{\sim}{A}=0.8/a+0.6/b+0.2/c,$$
$$\underset{\sim}{B}=0.5/d+0.9/e,$$

那么 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 是互斥的。事实上，可由

$$\underset{\sim}{A}(a)\wedge\underset{\sim}{B}(a)=0,\ \underset{\sim}{A}(b)\wedge\underset{\sim}{B}(b)=0,\ \underset{\sim}{A}(c)\wedge\underset{\sim}{B}(c)=0,$$
$$\underset{\sim}{A}(d)\wedge\underset{\sim}{B}(d)=0,\ \underset{\sim}{A}(e)\wedge\underset{\sim}{B}(e)=0$$

及定义 2.2.3 得出 $\underset{\sim}{A}$、$\underset{\sim}{B}$ 为互斥的 F 事件。

定义 2.2.4 若 $\underset{\sim}{A}_i$，$i=1,2,\cdots$为 F 事件，如果对任意的 $\omega\in\Omega$，有

$$\underset{\sim}{A}_i(\omega)\wedge\underset{\sim}{A}_j(\omega)=0,\ i\neq j,\ i,j=1,2,\cdots$$

则称$\{\underset{\sim}{A}_i, i=1,2,\cdots\}$两两互斥。

§3 公 理 结 构

经典概率论的基石是 Колмогоров 的公理，它给出了随

机事件与概率的公理化结构．我们目前的任务是研究 F 事件的概率，“概率”一词的意义是明确的，应遵从Колмогоров的公理，但是 F 事件却是经典事件概念的一般化，因此就必须改造 Колмогоров 公理．为了方便，我们先介绍Колмогоров 公理，然后将其推广至 F 集合．

设 Ω 为任一空间，$\mathscr{K}$ 为 Ω 的某些子集组成的集系，满足

1）$\Omega\in\mathscr{K}$；

2）$A_i\in\mathscr{K}$，$i=1,2,\cdots$ 则 $\bigcup\limits_i A_i\in\mathscr{K}$；

3）$A\in\mathscr{K}$ 则 $A^c\in\mathscr{K}$

则称 $\mathscr{K}$ 为 Ω 上的 σ 代数或完全可加系．又设 $P(A)$ 为对 $A\in\mathscr{K}$ 定义的集函数，满足

4）$0\leqslant P(A)\leqslant 1$；

5）$A_i\cap A_j=\phi$，$i\neq j$，$i,j=1,2,\cdots$，且 A_i，$A_j\in\mathscr{K}$ 则 $P\left(\bigcup\limits_i A_i\right)=\sum\limits_i P(A_i)$；

6）$P(\Omega)=1$，

如此的 $P(A)$ 称为 A 的概率，A 称为随机事件，$\Omega,\mathscr{K},P$ 一起 $(\Omega,\mathscr{K},P)$ 称为概率空间．

这就是Колмогоров 公理．对于 F 集合来说 Ω 上的 F 集系可构成 σ 软代数．由于 F 集合是经典集合的一般化，从而使我们考虑使用 σ 软代数，得到 F 事件概率的公理体系如下：

设 Ω 为任一空间，$\mathscr{F}(\Omega)$ 为 Ω 的 F 集合族，满足

(1) $1\in\mathscr{F}$；

(2) 若 $\underset{\sim}{A}_i\in\mathscr{F}$，$i=1,2,\cdots$ 则 $\bigvee\limits_i \underset{\sim}{A}_i\in\mathscr{F}$；

(3) 若 $\underset{\sim}{A}\in\mathscr{F}$，则 $\underset{\sim}{A}^c=1-\underset{\sim}{A}\in\mathscr{F}$．

又设映射 $\tilde{P}$：$\mathscr{F}\to[0,1]$，满足

(4) $\tilde{P}(1)=1$；

(5) $0\leqslant\tilde{P}(\underset{\sim}{A})\leqslant 1$；

(6) 若 $\underset{\sim}{A}_i\in\mathscr{F}$，$i=1,2,\cdots$，且 $\underset{\sim}{A}_i\wedge\underset{\sim}{A}_j=0$，$i\neq j$，$i,j=1,2,\cdots$，则

$$\tilde{P}\left(\bigvee_i \underset{\sim}{A}_i\right)=\sum_i\tilde{P}(\underset{\sim}{A}_i).$$

称 Ω 的元素为基本事件，$\underset{\sim}{A}$为 F 事件，$\mathscr{F}$为 F 事件域，$\tilde{P}$为定义在$\mathscr{F}$上的概率，$\tilde{P}(\underset{\sim}{A})$为 F 事件 $\underset{\sim}{A}$ 的概率。称$(\Omega,\mathscr{F},\tilde{P})$为不分明事件的概率空间，或 F 概率空间。

注意，这里$\tilde{P}(\underset{\sim}{A})$是由 F 子集 $\underset{\sim}{A}$ 的隶属函数$\underset{\sim}{A}(\cdot)$所确定的概率，即

$$\tilde{P}(\underset{\sim}{A}(u)),\quad \forall u\in\Omega$$

这个公理体系首先由 E. P. Klement 提出，后经多人修改，特别是 1982 年 P. Sments 的修改. 他提出用下面的(6′)代换(6)：

(6′)对一切 $\underset{\sim}{A},\underset{\sim}{B}\subseteq\Omega$，

$$\tilde{P}(\underset{\sim}{A}\vee\underset{\sim}{B})+\tilde{P}(\underset{\sim}{A}\wedge\underset{\sim}{B})=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B}),$$

但(6′)比(6)更强. 公理形式也有多种形式，如上面之(4)至(6)更改为下面之(Ⅳ)至(Ⅵ)：

(Ⅳ) $\tilde{P}(0)=0$ 且 $\tilde{P}(1)=1$；

(Ⅴ) $\tilde{P}(\underset{\sim}{A}\vee\underset{\sim}{B})+\tilde{P}(\underset{\sim}{A}\wedge\underset{\sim}{B})=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B})$；

(Ⅵ) 若$\underset{\sim}{A}_1\subset\underset{\sim}{A}_2\subset\cdots\subset\underset{\sim}{A}_n\subset\cdots$则

$$\tilde{P}\left(\bigvee_n \underset{\sim}{A}_n\right)=\sup_u\tilde{P}(\underset{\sim}{A}_n),$$

这里对 每个 $\alpha\in[0,1]$，$\underset{\sim}{\alpha}$ 表示一个 F 事件，对每一个$\omega\in\Omega$，

它的隶属函数值为 α。

当 $\underset{\sim}{A}=A$时，显然有

$$\tilde{P}(\underset{\sim}{A})=P(A).$$

例 2.3.1 设 $\Omega=\{a,b,c\}$,则 Ω 的一切子集之集$\mathscr{K}(\Omega)$共含有 2^3 个元素。即

$$\mathscr{K}(\Omega)=\{\phi,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\{b,c\},\Omega\}.$$

若指定

$$P(a)=0.2;\ P(b)=0.4;\ P(c)=0.4.$$

则由概率性质得出

$$\tilde{P}(\{a,b\})=0.6,$$

这是众所周知的事情。但是，在此若取隶属函数的值集为$M=\{0,0.5,0.8,1\}$时，则可得一切 F 子集族共有 $4^3=64$ 个F子集。假定 a,b,c 各自发生的概率仍是 0.2，0.4 及0.6。那么如何求出一个 F 事件的概率呢？如

$$\underset{\sim}{A}=0.5/a+0.8/b+1/c$$

怎么计算 $\tilde{P}(\underset{\sim}{A})$？虽然 0.5 及 0.8 是[0,1]中的数,但是它们是 a，b 隶属于集 $\underset{\sim}{A}$ 的程度，不是概率。

读者不难验证,$\mathscr{K}(\Omega)$是一σ事件域，而Ω 的一切F子集族$\mathscr{F}(\Omega)$是一σ软代数，这可妨照例 1.2.9的方法来验证。

§4 概率的赋值

利用概率的公理定义便于进行理论性的研究，但对一个具体问题却无法给出一种计算方法。这种具体的定义，我们称为概率的赋值法。早在 1968 年 Zadeh 给出 F 事件的概率定义，为了说清定义的实质，先从经典概率论中的随机变量的数学期望谈起。

随机事件 A 与经典集合 A 等同，当对经典集引入特征函数 I_A 后，I_A 与 A 等同．从而将特征函数 I_A 看作为一个取值为 0 与 1 的随机变量。

I_A	1	0
P	p	q

$p+q=1$，$p\geqslant 0$，$q\geqslant 0$。

于是有

$$EI_A=1\times p+0\times q=p=P(I_A=1).$$

即 $P(A)=EI_A$。

若设经典概率空间为$(\Omega,\mathscr{K},P)$．所谓 A 为一个随机事件是指 $A\in\mathscr{K}$，其上一个随机变量 ξ 是 Ω 上的一个映射 $\xi: \Omega\to R$，并且关于 $\mathscr{K}$ 可测，即对于直线 R 上任何一个$\alpha\in R$有

$$\{\omega\,|\,\xi(\omega)<\alpha\}\in\mathscr{K},$$

它等价于

$$\{\omega\,|\,\xi(\omega)\geqslant\alpha\}\in\mathscr{K}.$$

因此，一般地说，任何一个随机事件 A，

$$P(A)=EI_A=\int I_A(x)dP(x)$$

其中积分是 Lebesgue-Stielljes 积分。可见一个随机事件 A 的概率 $P(A)$是 A 的特征函数 I_A 的数学期望。按照这种思考方法，我们具体定义 F 事件 $\underset{\sim}{A}$ 的概率 $\underset{\sim}{P}(\underset{\sim}{A})$．

设$(\Omega,\mathscr{K},P)$为一经典概率空间，$\mathscr{K}$ 为一σ事件域。$\underset{\sim}{A}$为 Ω 上的一个 F 事件．而且 $\underset{\sim}{A}$ 的隶属函数 $\underset{\sim}{A}(\cdot): \Omega\to[0,1]$关于$\mathscr{K}$可测。

定义 2.4.1 F 事件 $\underset{\sim}{A}$ 的概率 $\underset{\sim}{P}(\underset{\sim}{A})$定义为

$$\underset{\sim}{P}(\underset{\sim}{A})=\int_\Omega \underset{\sim}{A}(x)dP(x).$$

当 $\Omega=\{x_1,x_2,\cdots\}$，且 $P(x_i)=p_i$，$i=1, 2, \cdots$，

$\sum_i p_i = 1$，$p_i \geq 0$ 时，则

$$\tilde{P}(\underset{\sim}{A}) = \sum_i \underset{\sim}{A}(x_i) p_i;$$

若 $\Omega = \{x\}$，F 事件 $\underset{\sim}{A}$ 的隶属函数 $\underset{\sim}{A}(x)$ 有概率密度函数 $f(x)$ 时，则

$$\tilde{P}(\underset{\sim}{A}) = \int_\Omega \underset{\sim}{A}(x) f(x)\, dx.$$

显然，当 $\underset{\sim}{A}$ 为非不分明时，$\tilde{P}(\underset{\sim}{A}) = P(\underset{\sim}{A})$。

特别我们可取 $\Omega = R^n$。此时，概率空间为 $(R^n.\mathscr{B},P)$，这里 $\mathscr{B}$ 是 R^n 的一切 Borel 集域，P 是概率. 称 R^n 上的一个 Borel 可测的 F 集合为 F 事件。因此 $\mathscr{P}_i(R^n)$(R^n 上一切 F 事件的全体)中的 F 事件 $\underset{\sim}{A}$ 由 $\underset{\sim}{A}(x)$: $R^n \to [0,1]$ 给定，如此若 $\mathscr{B}_{[0,1]}$ 是 $[0,1]$ 内 Borel 集族，$\underset{\sim}{A}^{-1}(\mathscr{B}_{[0,1]}) \in \mathscr{B}$. (在 $[0,1]$ 上的 Borel 集的逆像是 R^n 中的 Borel集)。更精确地说，若对任何 $\alpha \in [0,1]$ 有 $\{x \in R^n \mid \underset{\sim}{A}(x) \geq \alpha\} \in \mathscr{B}$. 至于 $\underset{\sim}{A}$ 的概率 $\tilde{P}(\underset{\sim}{A})$，应为

$$\tilde{P}(\underset{\sim}{A}) = \int_{R^n} \underset{\sim}{A}(x) dP(x).$$

上述概念的引入显得自然合理。我们在此仅对 $\Omega = R^n$ 的情形进行讨论，至于一般的 Ω 可根据 F集合的表现定理类似讨论。

定理 2.4.1 令 $\mathscr{F}_b(R^n)$ 是 R^n 中一切 F 事件之集，则 $(R^n, \mathscr{F}_b(R^n), \tilde{P})$ 是 F 概率空间。

证 我们必须先证明 $\mathscr{F}_b(R^n)$ 是一个 F 事件域。

设 $\underset{\sim}{A}$，$\underset{\sim}{B} \in \mathscr{F}_b(R^n)$，则 $\underset{\sim}{A} \vee \underset{\sim}{B}$，$\underset{\sim}{A} \wedge \underset{\sim}{B} \in \mathscr{F}_b(R^n)$，这是因为

$$\{x|(\underset{\sim}{A}\vee\underset{\sim}{B})(x)\geqslant\alpha\}=\{x|\underset{\sim}{A}(x)\geqslant\alpha\}\cup\{x|\underset{\sim}{B}(x)\geqslant\alpha\},$$

$$\{x|(\underset{\sim}{A}\wedge\underset{\sim}{B})(x)\geqslant\alpha\}=\{x|\underset{\sim}{A}(x)\geqslant\alpha\}\cap\{x|\underset{\sim}{B}(x)\geqslant\alpha\}.$$

另外，由于 R^n 是属于 $\mathscr{B}$ 的，所以 $1\in\mathscr{F}_b(R^n)$，且当 $\underset{\sim}{A}\in\mathscr{F}_b(R^n)$时。因为

$$\{x|(1-\underset{\sim}{A})(x)\geqslant\alpha\}=\{x|\underset{\sim}{A}(x)\leqslant1-\alpha\}\in\mathscr{B},$$

从而 $1-\underset{\sim}{A}\in\mathscr{F}_b(R^n)$。所以 $\mathscr{F}_b(R^n)$是 F 事件域。

现在我们来证明 $\tilde{P}$ 是一个 F 概率。

（1）因为 $P(R^n)=1$，所以 $\tilde{P}(1)=1$。

（2）因为$\underset{\sim}{A}\in\mathscr{F}_b(R^n)$ 时，$0\leqslant\underset{\sim}{A}(x)\leqslant1$，故$0\leqslant\tilde{P}(\underset{\sim}{A})\leqslant1$。

（3）若$\{\underset{\sim}{A}_i, i=1,2,\cdots\}\in\mathscr{F}_b(R^n)$，两两互斥，则

$$\begin{aligned}\tilde{P}\left(\bigvee_i\underset{\sim}{A}_i\right)&=\int_{R^n}\left(\bigvee_i\underset{\sim}{A}_i\right)(x)dP(x)\\&=\int_{\bigvee_i S_i}\sup\{\underset{\sim}{A}_i(x)\}dP(x)\\&=\sum_i\int_{S_i}\underset{\sim}{A}_i(x)dP(x)=\sum_i\tilde{P}(\underset{\sim}{A}_i)\end{aligned}$$

其中 $S_i=\{x|\underset{\sim}{A}_i(x)>0\}$，显然 $S_i\in R^n$，且 $S_i\cap S_j=\phi$。又因 $\bigvee_i S_i\subset R^n$，当 $x\in R^n-\bigvee_i S_i$时，$\underset{\sim}{A}_i(x)=0$.

由(1)-(3)证得 $\tilde{P}$ 是 $\mathscr{F}_b(R^n)$ 上的概率。从而 $(R^n,\mathscr{F}_b(R^n),\tilde{P})$为 F 概率空间。

为了明了(3)中的证明思路，我们用有限的 Ω 及两个 F 事件$\underset{\sim}{A}$ 及 $\underset{\sim}{B}$来说明。此时(3)成为：$\underset{\sim}{A}$、$\underset{\sim}{B}\in\mathscr{F}_b(\Omega)$，$\underset{\sim}{A}\cap\underset{\sim}{B}=\phi$，则

$$\tilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B})=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B}).$$

我们来证明它是正确的。因为

$$\tilde{P}(\underset{\sim}{A})=\sum_{i=1}^{n}\underset{\sim}{A}(x_i)p_i, \qquad \tilde{P}(\underset{\sim}{B})=\sum_{i=1}^{n}\underset{\sim}{B}(x_i)p_i$$

而

$$\tilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B})=\sum_{i=1}^{n}(\underset{\sim}{A}(x_i)\vee\underset{\sim}{B}(x_i))p_i$$

因为 $\underset{\sim}{A}\cap\underset{\sim}{B}=\phi$，即 $\underset{\sim}{A}(x_i)\wedge\underset{\sim}{B}(x_i)=0$，$x_i\in\Omega$，所以对任意 $x_i\in\Omega$ 有

$$\underset{\sim}{A}(x_i)\vee\underset{\sim}{B}(x_i)=\underset{\sim}{A}(x_i)+\underset{\sim}{B}(x_i)$$

由于

$$\underset{\sim}{A}(x_i)>0 \quad 则 \quad \underset{\sim}{B}(x_i)=0,$$
$$\underset{\sim}{B}(x_i)>0 \quad 则 \quad \underset{\sim}{A}(x_i)=0,$$

所以

$$\begin{aligned}\tilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B})&=\sum_{i=1}^{n}(\underset{\sim}{A}(x_i)+\underset{\sim}{B}(x_i))p_i\\&=\sum_{i=1}^{n}\underset{\sim}{A}(x_i)p_i+\sum_{i=1}^{n}\underset{\sim}{B}(x_i)p_i\\&=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B}).\end{aligned}$$

这个定理说明经典的概率空间$(R^n,\mathscr{B},P)$可扩张为F概率空间$(R^n,\mathscr{F}_b(R^n),\tilde{P})$，并且指明了概率 $\tilde{P}(\underset{\sim}{A})$赋值的合理性。

例 2.4.1 向一个目标进行射击直至击中为止。设各次射击是相互独立的，每次击中目标的概率为 p。但是，射击了不多几次就击中了目标，试求其概率。

解 设以 $\underset{\sim}{A}$ 表示“射击了不多几次就击中了目标”，且

$$\underset{\sim}{A}=1/1+0.8/2+0.6/3+0.4/4.$$

按照定义，$\underset{\sim}{A}$ 发生的概率为

$$\tilde{P}(\underset{\sim}{A})=\sum_{i=1}^{\infty}\underset{\sim}{A}(x_i)p_i=\sum_{i=1}^{\infty}\underset{\sim}{A}(x_i)(1-p)^{i-1}p$$
$$=p+0.8(1-p)p+0.6(1-p)^2p$$
$$+0.4(1-p)^3p。$$

例 2.4.2 已知某种产品的次品率为0.01，从中任取100个。以 $\underset{\sim}{A}$ 表示"所取的产品几乎无次品"，$\underset{\sim}{B}$表示"所取的产品大约有 4 个次品"：

$$\underset{\sim}{A}=1/0+1/1+0.8/2+0.5/3,$$
$$\underset{\sim}{B}=0.6/2+1/3+1/4+1/5+0.6/6,$$

试求 $\tilde{P}(\underset{\sim}{A})$及 $\tilde{P}(\underset{\sim}{B})$。

解 设 p_i 记 100 个产品中恰有 i 个次品的概率，

$$p_i=C_{100}^i(0.01)^i(1-0.01)^{100-i},$$
$$i=0，1，\cdots，100。$$

按 F 概率的定义，有

$$\tilde{P}(\underset{\sim}{A})=\underset{\sim}{A}(0)p_0+\underset{\sim}{A}(1)p_1+\underset{\sim}{A}(2)p_2+\underset{\sim}{A}(3)p_3$$
$$=0.37+0.37+0.8\times0.18+0.5\times0.06$$
$$=0.91,$$
$$\tilde{P}(\underset{\sim}{B})=B(0)p_0+\underset{\sim}{B}(1)p_1+\underset{\sim}{B}(2)p_2+\underset{\sim}{B}(3)p_3$$
$$+\underset{\sim}{B}(4)p_4+\underset{\sim}{B}(5)p_5+\underset{\sim}{B}(6)p_6$$
$$=0+0+0+0.6\times0.18+0.06+0.02+0.003$$
$$+0.6\times0.0005=0.21,$$

其中

$$p_0=0.37,\ p_1=0.37,\ p_2=0.18,\ p_3=0.06,$$
$$p_4=0.02,\ p_5=0.003\quad p_6=0.0005。$$

例 2.4.3 对某地区进行普查，得慢性气管炎的发病率如下：

年令分组	发病率(%)
15岁以下	2.8
15～25岁	3.7
25～35岁	5.0
35～45岁	7.6
45～55岁	11.0
55岁以上	14.2

设

$$\text{“青年”}=\frac{1}{(15\sim25)}+\frac{0.6}{(25\sim35)},$$

$$\text{“老年”}=\frac{1}{(55\text{岁以上})}+\frac{0.3}{(45\sim55)}.$$

分别求出该地区“青年”及“老年”人患慢性气管炎的概率。

解 “青年”人患慢性气管炎的概率为

$$1\times0.037+0.6\times0.05=0.067;$$

“老年”人患慢性气管炎的概率为

$$0.142\times1+0.3\times0.11=0.173.$$

例 2.4.4 某物体的长度为 a。现用一个仪器去测量，设无系统误差，测量的方差为 $\sigma^2(\sigma>0)$，且“测量的结果在 a 的附近”的隶属函数为

$$\underset{\sim}{A}(x)=e^{-\frac{(x-a)^2}{b}},$$

试求其概率。其中 $b>0$ 为一适当选择的参数。

解 根据概率统计的知识和题设，x 的分布密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

由 F 概率的定义知

$$\tilde{P}(A)=\int_{-\infty}^{\infty}\underset{\sim}{A}(x)f(x)dx$$

故

$$\tilde{P}(\underset{\sim}{A})=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^{-\frac{(x-a)^2}{b}}\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}dx$$

$$=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^{-\frac{(x-a)^2(2\sigma^2+b)}{2b\sigma^2}}dx$$

$$=\sqrt{\frac{b}{2\sigma^2+b}}\,.$$

例 2.4.5 设 $\underset{\sim}{A}\subset R^+$，其隶属函数为

$$\underset{\sim}{A}(x)=\begin{cases}1-\dfrac{x}{2}, & 0\leqslant x\leqslant 2,\\ 0\;, & 2<x.\end{cases}$$

而 R^+上概率分布函数为

$$F(x)=1-e^{-x},\quad x\in R^+.$$

求 $\tilde{P}(\underset{\sim}{A})$。

解 因为 $f(x)=F'(x)$，故

$$f(x)=\frac{dF(x)}{dx}=e^{-x},\quad x\geqslant 0.$$

由定义

$$\tilde{P}(\underset{\sim}{A})=\int_{R^+}\underset{\sim}{A}(x)dF(x)$$

$$=\int_0^2\left(1-\frac{x}{2}\right)e^{-x}dx$$

$$=\frac{1}{2}+\frac{e^{-2}}{2}=0.567.$$

例 2.4.6 设

$$E_1=\{a,b,c\},$$
$$E_2=\{\alpha,\beta,\gamma,\delta\}.$$

考虑 $E_1 \times E_2$ 上的不分明子集 $\underset{\sim}{G}$，其隶属函数见下表(a)。其在 $E_1 \times E_2$ 上的概率分布为 $P(x, y)$，如下表(b)所示。

表(a)

$\underset{\sim}{G}(x,y)$	α	β	γ	δ
a	0.8	1	0.2	0.4
b	0.9	0	0.5	0.5
c	0.1	0.8	0.6	0

表(b)

P	α	β	γ	δ
a	0	0.06	0.10	0.11
b	0.08	0	0.15	0.06
c	0.04	0.10	0.08	0.22

求 $\tilde{P}(\underset{\sim}{G})$。

解　由定义

$$
\begin{aligned}
\tilde{P}(\underset{\sim}{G}) &= \sum_x \sum_y \underset{\sim}{G}(x, y) \cdot P(x, y) \\
&= 0.8 \times 0 + 1 \times 0.06 + 0.2 \times 0.10 \\
&\quad + 0.4 \times 0.11 + 0.9 \times 0.08 + \\
&\quad 0 \times 0 + 0.5 \times 0.15 + 0.5 \times 0.06 \\
&\quad + 0.1 \times 0.04 + 0.8 \times 0.10 + 0.6 \\
&\quad \times 0.08 + 0 \times 0.22 \\
&= 0.469.
\end{aligned}
$$

由以上定义及诸例中可见，F 事件 $\underset{\sim}{A}$ 的概率 $\tilde{P}(\underset{\sim}{A})$表示 R^n具有不分明性质 $\underset{\sim}{A}$ 的程度，相应实验是对或多或少属于$\underset{\sim}{A}$的元素 x 的随机选择，每次试验给出一个隶属函数值$\underset{\sim}{A}(x)$。从而 $\underset{\sim}{A}(x)$反映关于事件 $\underset{\sim}{A}$ 的意义下的不分明性，$\tilde{P}(x)$是关于 x 的发生的随机性。因此，可以仿照经典概率的统计定义给$\tilde{P}(\underset{\sim}{A})$赋值为

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} \underset{\sim}{A}_j(x)
$$

其中 n 为实验的次数，$x \in R^n$。因而可以把 $\tilde{P}(\underset{\sim}{A})$ 解释成R^n中“属于”$\underset{\sim}{A}$的元素的浓度。

由于通常的交集、并集的运算在 F 集合可扩张为隶属函数的最小和最大运算。虽然 Bellman 和 Giertz证明了在某些适当的条件下，最大和最小运算是合适的（见张文修 $p.140$），但可不局限于此，有多种讨论方法，因此 F 事件的概率也将从各种不同的角度进行研究，特别值得注意的是从 F 测度来研究 F 概率的情形。

§5　F 事件的概率的性质

设经典概率空间为 $(R^n,\mathscr{B},P)$，F 概率空间为 $(R^n,\mathscr{F}_b(R^n),\widetilde{P})$，$\underset{\sim}{A}$、$\underset{\sim}{B}\in\mathscr{F}_b(R^n)$。

性质 1　$\widetilde{P}(0)=0$。

证　$\because \phi\in\mathscr{F}_b(b)$，　$\therefore \widetilde{P}(0)=0$。

性质 2　若 $\underset{\sim}{A}\subset\underset{\sim}{B}$ 则 $\widetilde{P}(\underset{\sim}{A})\leqslant\widetilde{P}(\underset{\sim}{B})$。

证　$\because \underset{\sim}{A}\subset\underset{\sim}{B}$　$\therefore \forall x\in R^n$，有 $\underset{\sim}{A}(x)\leqslant\underset{\sim}{B}(x)$，

$$\widetilde{P}(\underset{\sim}{A})=\int_{R^n}\underset{\sim}{A}(x)dP(x)\leqslant\int_{R^n}\underset{\sim}{B}(x)dP(x)=\widetilde{P}(\underset{\sim}{B})。$$

性质 3　$\widetilde{P}(\underset{\sim}{A}^c)=1-\widetilde{P}(\underset{\sim}{A})$.

证　$\because$　$\underset{\sim}{A}^c(x)=1-\underset{\sim}{A}(x)$，故

$$\widetilde{P}(\underset{\sim}{A}^c)=\int_{R^n}\underset{\sim}{A}^c(x)dP(x)=\int_{R^n}[1-\underset{\sim}{A}(x)]dP(x)$$

$$=1-\int_{R^n}\underset{\sim}{A}(x)dP(x)=1-\widetilde{P}(\underset{\sim}{A})。$$

性质 4　$\widetilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B})+\widetilde{P}(\underset{\sim}{A}\cap\underset{\sim}{B})=\widetilde{P}(\underset{\sim}{A})+\widetilde{P}(\underset{\sim}{B})$。

证　$\widetilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B})+\widetilde{P}(\underset{\sim}{A}\cap\underset{\sim}{B})$

$$=\int_{R^n}(\underset{\sim}{A}\cup\underset{\sim}{B})(x)dP(x)+\int_{R^n}(\underset{\sim}{A}\cap\underset{\sim}{B})(x)dP(x)$$

$$=\int_{R^n}(\underset{\sim}{A}(x)\vee\underset{\sim}{B}(x))dP(x)$$

$$+\int_{R^n}(\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x))dP(x)$$

$$=\int_{R^n}\max(\underset{\sim}{A}(x),\underset{\sim}{B}(x))dP(x)$$

$$+\int_{R^n}\min(\underset{\sim}{A}(x),\underset{\sim}{B}(x))dP(x)$$

$$=\int_{R^n}[\underset{\sim}{A}(x)+B(x)]dP(x)$$

$$=\int_{R^n}\underset{\sim}{A}(x)dP(x)+\int_{R^n}\underset{\sim}{B}(x)dP(x)$$

$$=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B}).$$

例 2.5.1 设 $R^n=\{x_1,x_2\}$，且

$$P(x_1)=P(x_2)=\frac{1}{2}.$$

$\underset{\sim}{A}=0.2/x_1+0.4/x_2$，$\underset{\sim}{B}=0.6/x_1+0.3/x_2$。

则

$$\tilde{P}(\underset{\sim}{A})=0.2\times\frac{1}{2}+0.4\times\frac{1}{2}=0.3;$$

$$\tilde{P}(\underset{\sim}{B})=0.6\times\frac{1}{2}+0.3\times\frac{1}{2}=0.45;$$

$$\tilde{P}(\underset{\sim}{A}\cap\underset{\sim}{B})=0.2\times\frac{1}{2}+0.3\times\frac{1}{2}=0.25;$$

$$\tilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B})=0.4\times\frac{1}{2}+0.6\times\frac{1}{2}=0.5.$$

故 $\tilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B})+\tilde{P}(\underset{\sim}{A}\cap\underset{\sim}{B})=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B})$。

性质 4 不难推广至多个 F 事件：

$$\begin{aligned}\tilde{P}(\underset{\sim}{A}_1+\underset{\sim}{A}_2+\underset{\sim}{A}_3)=&\tilde{P}(\underset{\sim}{A}_1)+\tilde{P}(\underset{\sim}{A}_2)+\tilde{P}(\underset{\sim}{A}_3)\\&-\tilde{P}(\underset{\sim}{A}_1\cap\underset{\sim}{A}_2)-\tilde{P}(\underset{\sim}{A}_2\cap\underset{\sim}{A}_3)\\&-\tilde{P}(\underset{\sim}{A}_1\cap\underset{\sim}{A}_3)\\&+\tilde{P}(\underset{\sim}{A}_1\cap\underset{\sim}{A}_2\cap A_3).\end{aligned}$$

一般地说

$$\widetilde{P}\left(\bigcup_{i=1}^{n} \underset{\sim}{A}_i\right)=\sum_{i=1}^{n}\widetilde{P}(\underset{\sim}{A}_i)-\sum_{i,j}\widetilde{P}(\underset{\sim}{A}_i\cap\underset{\sim}{A}_j)+\cdots$$

$$+(-1)^n\widetilde{P}\left(\bigcap_{i=1}^{n}\underset{\sim}{A}_i\right)$$

性质 5 设 $\underset{\sim}{A}_1$, $\underset{\sim}{A}_2$, …为 R^n 上的 F 事件，且$\underset{\sim}{A}_j\cap\underset{\sim}{A}_k=\phi$, $j\neq k$, $j,k=1, 2, \cdots$又$\bigcup_i \underset{\sim}{A}_i=R^n$，则$\sum_i \widetilde{P}(\underset{\sim}{A}_i)=1$。

证 对于任一 R^n 中的元素 x，总存在某个 i 有 $\underset{\sim}{A}_i(x)=1$，且对于任意 $j\neq i$ 使 $\underset{\sim}{A}_j(x)=0$。于是对 R^n的任一元素 x，有$\sum_i \underset{\sim}{A}_i(x)=1$。因 此

$$\begin{aligned}\sum_i \widetilde{P}(\underset{\sim}{A}_i)&=\sum_i\int_{R^n}\underset{\sim}{A}_i(x)dP(x)\\&=\int_{R^n}\sum_i A_i(x)dP(x)\\&=\widetilde{P}(R^n)=1.\end{aligned}$$

性质 6 设 $\underset{\sim}{A}_1$, $\underset{\sim}{A}_2$,…为 R^n 上的 F 事件，则

$$\widetilde{P}\left(\bigcup_i \underset{\sim}{A}_i\right)\leqslant\sum_i\widetilde{P}(\underset{\sim}{A}_i).$$

证 因为

$$\widetilde{P}\left(\bigcup_i \underset{\sim}{A}_i\right)=\int_{R^n}\bigvee_i\underset{\sim}{A}_i(x)dP(x)$$

$$\begin{aligned}\sum_i P(\underset{\sim}{A}_i)&=\sum_i\int_{R^n}\underset{\sim}{A}_i(x)dP(x)\\&=\int_{R^n}\sum_i\underset{\sim}{A}_i(x)dP(x)\end{aligned}$$

当 $a\in[0,1]$，$b\in[0,1]$时，$(a\vee b)\leqslant(a+b)$，所以

$$\bigvee_i \underset{\sim}{A}_i(x)\leqslant\sum_i \underset{\sim}{A}_i(x)$$

从而

$$\begin{aligned}\widetilde{P}\left(\bigcup_i \underset{\sim}{A}_i\right)&=\int_{R^n}\bigvee_i \underset{\sim}{A}_i(x)dP(x)\\&\leqslant\int_{R^n}\sum_i \underset{\sim}{A}_i(x)dP(x)\\&=\sum_i \widetilde{P}(\underset{\sim}{A}_i).\end{aligned}$$

注 F事件间运算关系中用隶属度之最大最小运算定义了并及交运算。但亦可将其考虑为代数和与代数积运算：

定义 2.5.1 在$\mathscr{F}(\Omega)$上定义$\oplus$和·运算，它们由

$$(\underset{\sim}{A}\oplus\underset{\sim}{B})(x)=\underset{\sim}{A}(x)+\underset{\sim}{B}(x)-\underset{\sim}{A}(x)\cdot\underset{\sim}{B}(x)$$

$$(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)=\underset{\sim}{A}(x)\cdot\underset{\sim}{B}(x)$$

分别确定，称为$\underset{\sim}{A}$，$\underset{\sim}{B}(\in\mathscr{F}(\Omega))$的代数和及代数积。

不难看出，对于$\{0,1\}$，运算$\oplus$和·分别与$\vee$和$\wedge$相同。

定义 2.5.2 在$(\Omega,\mathscr{F}(\Omega),\oplus,\cdot,c)$上一个$F$事件$\underset{\sim}{A}$的概率$\widetilde{P}(\underset{\sim}{A})$同定义 2.4.1 的一样。

在这个定义下，关于概率的这几个性质仍能成立。特别是

$$\widetilde{P}(\underset{\sim}{A}\oplus\underset{\sim}{B})=\widetilde{P}(\underset{\sim}{A})+\widetilde{P}(\underset{\sim}{B})-\widetilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B}).$$

事实上，由定义有

$$\begin{aligned}\widetilde{P}(\underset{\sim}{A}\oplus\underset{\sim}{B})&=\int_\Omega(\underset{\sim}{A}\oplus\underset{\sim}{B})(x)dP\\&=\int_\Omega[\underset{\sim}{A}(x)+\underset{\sim}{B}(x)-(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)]dP\\&=\int_\Omega\underset{\sim}{A}(x)dP+\int_\Omega\underset{\sim}{B}(x)dP\end{aligned}$$

$$-\int_{\Omega}(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)dP$$

$$=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B})-\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})$$

例 2.5.2 对于例 2.5.1 之$\underset{\sim}{A}$、$\underset{\sim}{B}$有

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=(0.2\times0.6)\times\frac{1}{2}+(0.4\times0.3)\times\frac{1}{2}$$

$$=\frac{1}{2}(0.12+0.12)=0.12,$$

$$\tilde{P}(\underset{\sim}{A}\oplus\underset{\sim}{B})=0.3+0.45-0.12=0.63。$$

直接由定义亦可得出

$$\tilde{P}(\underset{\sim}{A}\oplus\underset{\sim}{B})=[0.2+0.6-0.2\times0.6]\times\frac{1}{2}$$

$$+[0.3+0.4-0.3\times0.4]\times\frac{1}{2}$$

$$=0.63。$$

§6 F事件的广义概率

1982年 P。Sments将 Zadeh 的定义作了推广.

设 $\underset{\sim}{A}$ 是按照它的隶属函数 $\underset{\sim}{A}(x)$在非 F集$X=\{x\}$上定义的一个 F 集合。又设 $P(x)$是在 X 上定义的一个概率测度。

设符号+和∘表示语意并与或运算子，表征两个F 集合的并与交，它们由服从下面性质的 h 和 k 算子所给定：

P_1：$(\underset{\sim}{A}\circ\underset{\sim}{B})(x)=h(\underset{\sim}{A}(x),\ \underset{\sim}{B}(x))$；

P_2：$(\underset{\sim}{A}+\underset{\sim}{B})(x)=k(\underset{\sim}{A}(x),\ \underset{\sim}{B}(x))$；

P_3：$h(x,y)\in[0,1]$；

P_4：$k(x,y)\in[0,1]$；

P_5：$h(x,y)=h(y,x)$；

P_6：$(\underset{\sim}{A}\circ\underset{\sim}{B})\circ\underset{\sim}{C}=\underset{\sim}{A}\circ(\underset{\sim}{B}\circ\underset{\sim}{C})$；

P_7: $h(x,y)+k(x,y)=x+y$;

P_8: $h(x,y)\leqslant\min(x,y)$;

P_9: 若 $h(x,y)=0$ 则 $\min(x,y)=0$。

定理 2.6.1 $h(1,x)=x$.

证 若 $h(1,x)<x$，由 P_7 得 $k(1,x)>1$，这与 P_4 相矛盾，得证 $h(1,x)=x$。

读者不难验证，当取 + 和 ∘ 分别为∨与∧运算时，则

$$h(x,y)=x\wedge y;\ \ k(x,y)=x\vee y.$$

而当分别取为代数和⊕和代数积∘运算时，则

$$h(x,y)=xy;\ \ k(x,y)=x+y-xy.$$

设函数 g：$[0,1]\to R$，单调不减且 $g(0)=0, g(1)=1$。

定义 2.6.1 F事件 $\underset{\sim}{A}$ 的概率

$$\widetilde{P}(\underset{\sim}{A})=\int_X g(\underset{\sim}{A}(x))dP(x).$$

显然 Zadeh 的 F 事件 $\underset{\sim}{A}$ 的概率为其特殊情形。

定理 2.6.2 $(X,\mathscr{R}_b(x),\widetilde{P})$为 F 概率空间。

证 只要验证 $\widetilde{P}(\underset{\sim}{A})$ 满足 kolomogoroff 公理即可。显然

1) $0\leqslant\widetilde{P}(\underset{\sim}{A})\leqslant1$; 2) $\widetilde{P}(1)=1$。

3) 若 $\underset{\sim}{A}\subset X$，$\underset{\sim}{B}\subset X$，且 $\underset{\sim}{A}\circ\underset{\sim}{B}=\{\phi\}$，则

$$\widetilde{P}(\underset{\sim}{A}+\underset{\sim}{B})=\widetilde{P}(\underset{\sim}{A})+\widetilde{P}(\underset{\sim}{B}).$$

事实上，因为 $\underset{\sim}{A}\circ\underset{\sim}{B}=\{\phi\}$，则对任何 $x\in X$ 有

$$(\underset{\sim}{A}\circ\underset{\sim}{B})(x)=h(\underset{\sim}{A}(x),\underset{\sim}{B}(x))=0,$$

及$\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x)=0$(由 P_9)。设 $X_{\underset{\sim}{A}}=\{x|\ \underset{\sim}{A}(x)=0\}$，$X_{\underset{\sim}{B}}=\{x|\ \underset{\sim}{B}(x)=0\}$，于是 $X=X_{\underset{\sim}{A}}+X_{\underset{\sim}{B}}$。对任何 $x\in X_{\underset{\sim}{A}}\cap X_{\underset{\sim}{B}}$，有$(\underset{\sim}{A}+\underset{\sim}{B})(x)=0$(由$P_7,P_8$)。如此 $g[(\underset{\sim}{A}+\underset{\sim}{B})(x)]=0$。当$x\in X_{\underset{\sim}{A}}$时，

$(\underset{\sim}{A}+\underset{\sim}{B})(x)=\underset{\sim}{B}(x)$；当 $x\in X_{\underset{\sim}{B}}$时，$(\underset{\sim}{A}+\underset{\sim}{B})(x)=\underset{\sim}{A}(x)$. 于是

$$\tilde{P}(\underset{\sim}{A})=\int_X g(\underset{\sim}{A}(x))dP(x)=\int_{X_{\underset{\sim}{B}}} g[\underset{\sim}{A}(x)dP(x)$$

$$=\int_{X_{\underset{\sim}{B}}} g[(\underset{\sim}{A}+\underset{\sim}{B})(x)]dP(x).$$

如果 $\underset{\sim}{A}(x)>0$，只要假设 $\underset{\sim}{B}(x)=0$，即 $x\in X_{\underset{\sim}{B}}$. 类似可得

$$\tilde{P}(\underset{\sim}{B})=\int_{X_{\underset{\sim}{A}}} g[(\underset{\sim}{A}+\underset{\sim}{B})(x)]dP(x)$$

故

$$\tilde{P}(\underset{\sim}{A}+\underset{\sim}{B})=\left\{\int_{X_{\underset{\sim}{A}}}+\int_{X_{\underset{\sim}{B}}}\right\}g[(\underset{\sim}{A}+\underset{\sim}{B})(x)]dP(x)$$

$$=\tilde{P}(\underset{\sim}{A})+\tilde{P}(\underset{\sim}{B}).$$

定理得证。

§7 条件概率

2.7.1 F事件的条件概率

事件的条件概率可以像经典的条件概率一样来定义。并且仍用记号 $\underset{\sim}{A}|\underset{\sim}{B}$表示在已知 $\underset{\sim}{B}$发生条件下 $\underset{\sim}{A}$ 发生这样的 F 事件。在定义 1.6.1 的基础上定义 $\underset{\sim}{A}|\underset{\sim}{B}$ 的概率。

定义 2.7.1 当 $\tilde{P}(\underset{\sim}{B})>0$ 时，比值

$$\frac{\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})}{\tilde{P}(\underset{\sim}{B})}$$

称为在已知 F 事件 $\underset{\sim}{B}$ 发生条件下 F 事件 $\underset{\sim}{A}$ 发生的概率，记作 $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})$。

为了说明定义的合理性，我们首先证一个引理。

引理 若 $\underset{\sim}{A}$、$\underset{\sim}{B}$ 非空，则$\underset{\sim}{A}\cap\underset{\sim}{B}=\phi$ 与$\underset{\sim}{A}\cdot\underset{\sim}{B}=\phi$ 等价。

证 先证 $\underset{\sim}{A}\cap\underset{\sim}{B}=\phi\Rightarrow\underset{\sim}{A}\cdot\underset{\sim}{B}=\phi$。因为

$$(\underset{\sim}{A}\cap\underset{\sim}{B})(x)=\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x)=0,$$

当 $\underset{\sim}{A}(x)>0$ 时，则 $\underset{\sim}{B}(x)=0$；

当 $\underset{\sim}{B}(x)>0$ 时，则 $\underset{\sim}{A}(x)=0$。

所以　$\underset{\sim}{A}(x)\cdot\underset{\sim}{B}(x)=0$，即 $\underset{\sim}{A}\cdot\underset{\sim}{B}=\phi$。

再证　$\underset{\sim}{A}\cdot\underset{\sim}{B}=\phi\Rightarrow\underset{\sim}{A}\cap\underset{\sim}{B}=\phi$。

因为　　$\underset{\sim}{A}\cdot\underset{\sim}{B}=\phi$，有

$$\underset{\sim}{A}(x)\cdot\underset{\sim}{B}(x)=0,$$

此时 $\underset{\sim}{A}(x)$ 与 $\underset{\sim}{B}(x)$ 中至少有一个为零。若 $\underset{\sim}{A}(x)=0$，则 $\underset{\sim}{B}(x)\geqslant0$；若 $\underset{\sim}{B}(x)=0$，则 $\underset{\sim}{A}(x)\geqslant0$。从而

$$\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x)=0,$$

即 $\underset{\sim}{A}\cap\underset{\sim}{B}=\phi$。

定理2.7.1　若 $\underset{\sim}{A}\cap\underset{\sim}{B}=\phi$，则

$$\tilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B}|\underset{\sim}{C})=\tilde{P}(\underset{\sim}{A}|\underset{\sim}{C})+\tilde{P}(\underset{\sim}{B}|\underset{\sim}{C})$$

证　因为 $\underset{\sim}{A}\cap\underset{\sim}{B}=\phi$，所以 $\underset{\sim}{A}\cdot\underset{\sim}{B}=\phi$。于是

$$(\underset{\sim}{A}\cdot\underset{\sim}{C})\cdot(\underset{\sim}{B}\cdot\underset{\sim}{C})=\phi.$$

从而

$$\tilde{P}[(\underset{\sim}{A}\cdot\underset{\sim}{C})\cup(\underset{\sim}{B}\cdot\underset{\sim}{C})]=\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{C})+\tilde{P}(\underset{\sim}{B}\cdot\underset{\sim}{C})$$

但是

$$(\underset{\sim}{A}\cdot\underset{\sim}{C})\cup(\underset{\sim}{B}\cdot\underset{\sim}{C})=(\underset{\sim}{A}\cup\underset{\sim}{B})\cdot\underset{\sim}{C}$$

所以

$$\tilde{P}[(\underset{\sim}{A}\cup\underset{\sim}{B})\cdot\underset{\sim}{C}]=\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{C})+\tilde{P}(\underset{\sim}{B}\cdot\underset{\sim}{C})$$

由于 $\tilde{P}(\underset{\sim}{C})>0$，故等式两端同除 $\tilde{P}(\underset{\sim}{C})$ 得

$$\frac{\tilde{P}[(\underset{\sim}{A}\cup\underset{\sim}{B})\cdot C]}{\tilde{P}(\underset{\sim}{C})}=\frac{\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{C})}{\tilde{P}(\underset{\sim}{C})}=\frac{\tilde{P}(\underset{\sim}{B}\cdot\underset{\sim}{C})}{\tilde{P}(\underset{\sim}{C})}$$

即　$\tilde{P}(\underset{\sim}{A}\cup\underset{\sim}{B}|\underset{\sim}{C})=\tilde{P}(\underset{\sim}{A}|\underset{\sim}{C})+\tilde{P}(\underset{\sim}{B}|\underset{\sim}{C})$。

定理 2.7.2　$\tilde{P}(\Omega|\underset{\sim}{C})=1$

证　因为

$$\tilde{P}(\Omega|\underset{\sim}{C})=\frac{\tilde{P}(\Omega\cdot\underset{\sim}{C})}{\tilde{P}(\underset{\sim}{C})}=\frac{\tilde{P}(\underset{\sim}{C})}{\tilde{P}(\underset{\sim}{C})}=1.$$

定理 2.7.3 $0\leqslant\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})\leqslant 1$.

证 因为 $0\leqslant\underset{\sim}{A}(x)\leqslant 1$，所以

$$0\leqslant(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)=\underset{\sim}{A}(x)\times\underset{\sim}{B}(x)\leqslant\underset{\sim}{B}(x).$$

从而

$$0\leqslant\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})\leqslant\tilde{P}(\underset{\sim}{B})$$

两端同除 $\tilde{P}(\underset{\sim}{B})$ 得 $0\leqslant P(\underset{\sim}{A}|\underset{\sim}{B})\leqslant 1$.

从以上三个定理可以看出 F 事件的条件概率满足 Kolmogoroff 的公理，从而说明了定义的合理性．下面我们用例子来说明如何计算条件概率.

例 2.7.1 设

$$\Omega=\{x_1,x_2,x_3,x_4,x_5\}$$

$$P(x_1)=0.1,\quad P(x_2)=0.3,\quad P(x_3)=0.4,$$

$$P(x_4)=0,\qquad P(x_5)=0.2.$$

考虑

$$\underset{\sim}{A}=0.8/x_1+1/x_2+0.2/x_3+0.9/x_4,$$

$$\underset{\sim}{B}=1/x_1+0.4/x_2+0.6/x_3+0.7/x_4+0.8/x_5$$

于是

$$(\underset{\sim}{A}\cdot\underset{\sim}{B})=0.8/x_1+0.4/x_2+0.12/x_3+0.63/x_4+0/x_5$$

根据概率的定义很容易算出

$$\tilde{P}(\underset{\sim}{A})=0.46,\quad\tilde{P}(\underset{\sim}{B})=0.62,\quad\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=0.248$$

再由条件概率定义有

$$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\frac{\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})}{\tilde{P}(\underset{\sim}{B})}=\frac{0.248}{0.62}=0.4$$

同理可得

$$\tilde{P}(\underset{\sim}{B}|\underset{\sim}{A})=\frac{\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})}{\tilde{P}(\underset{\sim}{A})}=\frac{0.248}{0.46}=0.539$$

定义 2.7.1 中的 $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})$亦可写成

$$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\frac{\int_{\Omega}(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)dF(x)}{\int_{\Omega}\underset{\sim}{B}(x)dF(x)}$$

其中 $F(x)$为定义在 R 上的分布函数．这个公式运用起来 比较方便．

例 2.7.2 设 $\Omega=R^{+}=[0,+\infty)$，其上两个 F事件$\underset{\sim}{A}$，$\underset{\sim}{B}$：

$$\underset{\sim}{A}(x)=\begin{cases}1-\dfrac{x}{2}, & 0\leqslant x\leqslant 2,\\ 0, & 2<x.\end{cases}$$

$$\underset{\sim}{B}(x)=\begin{cases}\dfrac{x}{2}, & 0\leqslant x\leqslant 2,\\ 0 & 2<x.\end{cases}$$

又设在 R^{+}上定义了概率密度

$$f(x)=\frac{dF(x)}{dx}=e^{-x},\quad x\geqslant 0.$$

现在来求 $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})$及 $\tilde{P}(\underset{\sim}{B}|\underset{\sim}{A})$．因为

$$(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)=\underset{\sim}{A}(x)\cdot\underset{\sim}{B}(x)$$

$$=\begin{cases}\dfrac{x}{2}\left(1-\dfrac{x}{2}\right), & 0\leqslant x\leqslant 2,\\ 0, & 2<x\end{cases}$$

故有

$$\tilde{P}(\underset{\sim}{A})=\int_{R^{+}}\underset{\sim}{A}(x)dF(x)=\int_{0}^{2}\left(1-\frac{x}{2}\right)e^{-x}dx+\int_{2}^{\infty}0\cdot e^{-x}dx$$

$$=\int_{0}^{2}e^{-x}dx-\int_{0}^{2}\frac{x}{2}e^{-x}dx$$

$$= -e^{-x}\Big|_0^2 - \frac{1}{2}e^{-x}(x+1)\Big|_0^2$$

$$= -e^{-2}+1+\frac{1}{2}[3e^{-2}-1]$$

$$= -e^{-2}+\frac{3}{2}e^{-2}+\frac{1}{2}=\frac{1}{2}[e^{-2}+1]$$

$$=0.567;$$

$$\tilde{P}(\underset{\sim}{B})=0.296;$$

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=\int_{R^+}(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)dF(x)$$

$$=\int_0^2\frac{x}{2}\left(1-\frac{x}{2}\right)e^{-x}dx+\int_2^\infty 0\cdot e^{-x}dx$$

$$=\int_0^2\frac{x}{2}e^{-x}dx-\int_0^2\frac{x^2}{4}e^{-x}dx$$

$$=\left\{\frac{1}{2}e^{-x}(-x-1)+\frac{1}{4}e^{-x}(x^2+2x+2)\right\}\Big|_0^2$$

$$=e^{-2}\left[\frac{1}{2}(-3)+\frac{1}{4}\times 10\right]+\frac{1}{2}-\frac{1}{4}\times 2$$

$$=e^{-2}=0.136$$

所以

$$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\frac{P(\underset{\sim}{A}\cdot\underset{\sim}{B})}{P(\underset{\sim}{B})}=\frac{0.136}{0.296}=0.46,$$

$$\tilde{P}(\underset{\sim}{B}|\underset{\sim}{A})=0.240.$$

在条件概率的定义中，我们考虑了 F 事件 $(\underset{\sim}{A}\cap\underset{\sim}{B})$ 为 $(\underset{\sim}{A}\cdot\underset{\sim}{B})$。在普通集合时有

$$A\subset B\Rightarrow A\cap B=A$$

但对于 F 集合，则

由 $\underset{\sim}{A}\subset\underset{\sim}{B}$ 推不出 $\underset{\sim}{A}\cdot\underset{\sim}{B}=\underset{\sim}{A}$。

当 $\underset{\sim}{A}\cdot\underset{\sim}{B}=\phi$ 时，显然有 $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=0$。

2.7.2 Smets 的条件概率

当我们考虑 Smets 的广义概率时，条件概率的定义仍同上面一样。若 $g(u)$是单调不减函数，则有

定理 2.7.4 $0\leqslant\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})\leqslant1$.

证 由 P_8 知对任意的$x\in\Omega$，有$(\underset{\sim}{A}\circ\underset{\sim}{B})(x)\leqslant\underset{\sim}{B}(x)$。因为

$$0\leqslant g[(\underset{\sim}{A}\circ\underset{\sim}{B})(x)]\leqslant g[\underset{\sim}{B}(x)]$$

$$0\leqslant\tilde{P}(\underset{\sim}{A}\circ\underset{\sim}{B})\leqslant\tilde{P}(\underset{\sim}{B})$$

因为 $\tilde{P}(\underset{\sim}{A})>0$，所以

$$0\leqslant\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})\leqslant\tilde{P}(\underset{\sim}{B}).$$

定理 2.7.5 $\tilde{P}(\Omega|\underset{\sim}{B})=1$.

证 对于任意的$x\in\Omega$，有 $\Omega(x)=1$。由 P_{10} 有

$$(\Omega\circ\underset{\sim}{B})(x)=\underset{\sim}{B}(x)$$

所以

$$\tilde{P}(\Omega\circ\underset{\sim}{B})=\tilde{P}(\underset{\sim}{B})$$

于是

$$\tilde{P}(\Omega|\underset{\sim}{B})=1.$$

定理 2.7.6 令 $\underset{\sim}{A}$、$\underset{\sim}{B}$、$\underset{\sim}{C}$ 为 Ω 上三个 F 事件，$\underset{\sim}{A}\circ\underset{\sim}{B}\circ\underset{\sim}{C}=\{\phi\}$，则

$$\tilde{P}[(\underset{\sim}{A}+\underset{\sim}{B})|\underset{\sim}{C}]=\tilde{P}(\underset{\sim}{A}|\underset{\sim}{C})+\tilde{P}(\underset{\sim}{B}|\underset{\sim}{C}).$$

证 限制于空间 $\underset{\sim}{C}=\{x:\underset{\sim}{C}(x)>0\}$之上，应用可加性公理即可得证。

需要注意，若 $h(x,y)$等幂，正如同没有代数积最小运算子的情况一样，$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{A})=1$.

对于条件概率 $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})$，按照广义概率的定义，具有如下的表现：

$$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\int_{\Omega}g[(\underset{\sim}{A}|\underset{\sim}{B})(x)]dP(x)$$

其中

$$g[(\underset{\sim}{A}|\underset{\sim}{B})(x)]=g[(\underset{\sim}{A}\circ\underset{\sim}{B})(x)]/g[\underset{\sim}{B}(x)],$$

且

$$dP(x|\underset{\sim}{B})=g[\underset{\sim}{B}(x)]dP(x)/\widetilde{P}(\underset{\sim}{B}),$$

而函数 g 是一个单调不减函数，$g[(\underset{\sim}{A}\circ\underset{\sim}{B})(x)]\in[0,1]$，且 $dP(x)$是一个真正的概率测度。

若 $g(u)$是一个一一对应函数，当 $g(u)=u$ 时，能够定义一个条件隶属函数

$$(\underset{\sim}{A}|\underset{\sim}{B})(x)=g^{-1}(g[(\underset{\sim}{A}\circ\underset{\sim}{B})(x)]/g(\underset{\sim}{B}(x))).$$

表示在 x 属于 $\underset{\sim}{B}$ 的情况下 x 关于 $\underset{\sim}{A}$ 的隶属函数，记为

$$(\underset{\sim}{A}|\underset{\sim}{B})(x)=(\underset{\sim}{A}\circ\underset{\sim}{B})(x)/\underset{\sim}{B}(x).$$

由此可以得出结论，任何单调不减函数可以用定义 2.6.1 来定义一个 F 事件的概率。

2.7.3 乘法公式与独立性

不分明事件的独立性概念完全仿照通常的独立性概念来定义。从而使得由条件概率推导出的许多结论，对于不分明事件的情形照样成立。

定义 2.7.2 若 $\widetilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\widetilde{P}(\underset{\sim}{A})$，则称 $\underset{\sim}{A}$ 和 $\underset{\sim}{B}$ 是独立的。

注 如果我们按照通常的独立性定义那样，定义 $\underset{\sim}{A}$ 对于 $\underset{\sim}{B}$ 独立为使 $\widetilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\widetilde{P}(\underset{\sim}{A})$成立。那么，显然可定义 $\underset{\sim}{B}$ 对于 $\underset{\sim}{A}$ 独立，从而，可定义 $\underset{\sim}{A}$ 与$\underset{\sim}{B}$ 相互独立的概念。

定理 2.7.7 若 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 独立，则

$$\widetilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=\widetilde{P}(\underset{\sim}{A})\widetilde{P}(\underset{\sim}{B}).\qquad(1)$$

证 因为

$$\widetilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\frac{\widetilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})}{\widetilde{P}(\underset{\sim}{B})},\quad P(\underset{\sim}{B})>0,$$

而且由题设 $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{B})=\tilde{P}(\underset{\sim}{A})$。代入上式即得证。

定理 2.7.8 若$\tilde{P}(\underset{\sim}{B})>0$，$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=\tilde{P}(\underset{\sim}{A})\tilde{P}(\underset{\sim}{B})$，则 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 独立。

证明是显然的。

由这两个定理可见，若$(\underset{\sim}{A}\cdot\underset{\sim}{B})$的概率等于 $\underset{\sim}{A}$ 的概率与 $\underset{\sim}{B}$ 的概率之积，亦可作为 $\underset{\sim}{A}$ 与$\underset{\sim}{B}$ 独立的定义。另外，像通常概率乘法公式一样，称式(1)为不分明事件概率的乘法公式.

例 2.7.3 设 $X=\{x_1,x_2,x_3,x_4,x_5\}$,

$$P(x_1)=\frac{1}{60},\quad P(x_2)=\frac{1}{10},\quad P(x_3)=\frac{1}{3},$$

$$P(x_4)=\frac{1}{20},\quad P(x_5)=\frac{1}{2};$$

$$\underset{\sim}{A}=1/x_1+1/x_2+0/x_3+1/x_4+0/x_5;$$

$$\underset{\sim}{B}=1/x_1+0/x_2+1/x_3+1/x_4+0/x_5.$$

试判断 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 独立。

解 由题设可得

$$(\underset{\sim}{A}\cdot\underset{\sim}{B})=1/x_1+1/x_4,$$

所以

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=\frac{1}{60}+\frac{3}{60}=\frac{2}{30},$$

$$\tilde{P}(\underset{\sim}{A})=\frac{1}{60}+\frac{1}{10}+\frac{1}{20}=\frac{1}{6};$$

$$\tilde{P}(\underset{\sim}{B})=\frac{1}{60}+\frac{1}{3}+\frac{1}{20}=\frac{2}{5}.$$

可见

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=\tilde{P}(\underset{\sim}{A})\tilde{P}(\underset{\sim}{B}).$$

故 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 独立。

例 2.7.4 设 $X=\{x_1,x_2,x_3\}$,

$$P(x_1)=0.2,\ P(x_2)=0.7,\ P(x_3)=0.1,$$

$$\underset{\sim}{A}=0.680/x_1+1/x_2+0.130/x_3,$$

$$\underset{\sim}{B}=0.240/x_1+0.350/x_2+0.401/x_3.$$

我们来判断 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 是否独立。

由题设可得

$$(\underset{\sim}{A}\cdot\underset{\sim}{B})=0.16320/x_1+0.35000/x_2+0.05213/x_3.$$

且由不分明事件概率可得

$$\widetilde{P}(\underset{\sim}{A})=0.849,\ \widetilde{P}(\underset{\sim}{B})=0.3331,$$

$$\widetilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=0.2828.$$

从而得到 $\widetilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=\widetilde{P}(\underset{\sim}{A})\widetilde{P}(\underset{\sim}{B})$，故 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 独立。

例 2.7.5 设上例中的条件改为

$$P(x_1)=0.5,\ P(x_2)=0.1,\ P(x_3)=0.4,$$

考虑 $\underset{\sim}{A}$ 及 $\underset{\sim}{B}$，则有

$$\widetilde{P}(\underset{\sim}{A})=0.4920,\ \widetilde{P}(\underset{\sim}{B})=0.3154,$$

$$\widetilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=0.1374,$$

于是

$$\widetilde{P}(\underset{\sim}{A})\widetilde{P}(\underset{\sim}{B})=0.1552\neq 0.1374=\widetilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{B}).$$

所以 $\underset{\sim}{A}$ 与 $\underset{\sim}{B}$ 不独立。

顺便指出，在条件概率及独立性定义中使用 $(\underset{\sim}{A}\cdot\underset{\sim}{B})$，而不用$\underset{\sim}{A}\cap\underset{\sim}{B}$。原因在于采用$P(\underset{\sim}{A}\cap\underset{\sim}{B})$会得出错误的结论。如在例 2.7.4 中，我们考虑到

$$(\underset{\sim}{A}\cap\underset{\sim}{B})(x)=\underset{\sim}{A}(x)\wedge\underset{\sim}{B}(x)=0.24/x_1+0.35/x_2+0.130/x_3.$$

就有

$$\widetilde{P}(\underset{\sim}{A}\cap\underset{\sim}{B})=0.24\times0.2+0.35\times0.7+0.13\times0.1=0.306,$$

它不等于 0.2828. 从而

$$\underset{\sim}{P}(\underset{\sim}{A}\cap\underset{\sim}{B})\neq\underset{\sim}{P}(\underset{\sim}{A}\cdot\underset{\sim}{B})=\underset{\sim}{P}(\underset{\sim}{A})\underset{\sim}{P}(\underset{\sim}{B}).$$

其原因在于，一般说来，$(\underset{\sim}{A}\cap\underset{\sim}{B})(x)\neq(\underset{\sim}{A}\cdot\underset{\sim}{B})(x)$.

若取 $X_1=R^n$，$X_2=R^n$ 且 P 是乘积测度 $P_1\times P_2$，其中 P_1 和 P_2 分别是 X_1 和 X_2 上的概率测度. 又设 $\underset{\sim}{A}_1$ 和 $\underset{\sim}{A}_2$ 分别是 X_1 和 X_2 上的 F 事件，其特征用隶属函数 $\underset{\sim}{A}_1(x_1,x_2)=\underset{\sim}{A}_1(x_1)$ 和 $\underset{\sim}{A}_2(x_1,x_2)=\underset{\sim}{A}_2(x_2)$ 分别表示. 按照我们上面定义 $\underset{\sim}{A}_1$ 和 $\underset{\sim}{A}_2$ 是独立的，但若用 $P(\underset{\sim}{A}\cap\underset{\sim}{B})$ 来定义独立性则不行.

§8 Bayes公式及全概率公式

我们可以像通常概率论一样，讨论 Bayes 公式. 回忆那时的 Bayes 公式及全概率公式可叙述为：

设 $E_1\cup E_2\cup\cdots\cup E_n=E.$ (2.8.1)

$E_i\cap E_j=\phi,\quad i\neq j,\quad i,j=1,2,\cdots,n$ (2.8.2)

$A\subset E,\ A=(A\cap E_1)\cup(A\cap E_2)\cup\cdots\cup(A\cap E_n)$ (2.8.3)

且已知

$P(E_i)>0,\qquad i=1,2,\cdots,n.$ (2.8.4)

$P(A|E_i),\qquad i=1,2,\cdots,n.$ (2.8.5)

那么有全概率公式

$$P(A)=\sum_{i=1}^{n}P(A|E_i)P(E_i) \tag{2.8.6}$$

及 Bayes 公式

$$P(E_i|A)=\frac{P(A|E_i)P(E_i)}{\sum_{i=1}^{n}P(A|E_i)P(E_i)},\quad i=1,2,\cdots,n \tag{2.8.7}$$

这些我们比较熟悉. 但是对于不分明事件 Bayes 公式应是什么样子呢?

设我们的论域为 E。$E_1,E_2,\cdots,E_n$为其上的 F 子集，且设

$$\underset{\sim}{E}_1\cup\underset{\sim}{E}_2\cup\cdots\cup\underset{\sim}{E}_n=E, \tag{2.8.8}$$

$$\underset{\sim}{A}\subset E,\quad E_i\subset E,\quad i=1,2,\cdots,n \tag{2.8.9}$$

$$E=\{x_1,x_2,\cdots,x_m\},m\text{ 为一有限数} \tag{2.8.10}$$

又设

$$\underset{\sim}{E}_i(x),\quad i=1,2,\cdots,n \tag{2.8.11}$$

为 $\underset{\sim}{E}_i$的隶属函数。

另设 $\underset{\sim}{A}\cdot\underset{\sim}{E}_i\subset\underset{\sim}{A},\quad i=1,2,\cdots,n.$ (2.8.12)

$$(\underset{\sim}{A}\cdot\underset{\sim}{E}_1)\cup(\underset{\sim}{A}\cdot\underset{\sim}{E}_2)\cup\cdots\cup(\underset{\sim}{A}\cdot\underset{\sim}{E}_n)=\underset{\sim}{A}\cdot E=A, \tag{2.8.13}$$

由于

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_i)=\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_i)\tilde{P}(\underset{\sim}{E}_i),\quad i=1,2,\cdots,n.$$

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_i)=\tilde{P}(\underset{\sim}{E}_i|\underset{\sim}{A})\tilde{P}(\underset{\sim}{A}),\quad i=1,2,\cdots,n.$$

于是

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j)=\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j)\tilde{P}(\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j),$$
$$i,j=1,2,\cdots,n,i<j.$$

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j\cdot\underset{\sim}{E}_k)=\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j\cdot\underset{\sim}{E}_k)$$
$$\times P(\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j\cdot\underset{\sim}{E}_k)$$
$$i,j,k=1,2,\cdots,n,\quad i<j<k.$$

……

$$\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\cdots\underset{\sim}{E}_n)$$
$$=\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\cdots\cdot\underset{\sim}{E}_n)\tilde{P}(\underset{\sim}{E}_1\cdot\cdots\cdot\underset{\sim}{E}_n) \tag{2.8.14}$$

从而

$$
\begin{aligned}
\tilde{P}(\underset{\sim}{A}) &= \tilde{P}[(\underset{\sim}{A}\cdot\underset{\sim}{E}_1)\cup(\underset{\sim}{A}\cdot\underset{\sim}{E}_2)\cup\cdots\cup(\underset{\sim}{A}\cdot\underset{\sim}{E}_n)] \\
&= \tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_1)+\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_2)+\cdots+\tilde{P}(\underset{\sim}{A}\cdot E_n) \\
&\quad -\tilde{P}[(\underset{\sim}{A}\cdot\underset{\sim}{E}_1)\cup(\underset{\sim}{A}\cdot\underset{\sim}{E}_2)] \\
&\quad -\tilde{P}[(\underset{\sim}{A}\cdot\underset{\sim}{E}_1)\cup(\underset{\sim}{A}\cdot\underset{\sim}{E}_3)] \\
&\quad -\cdots-\tilde{P}[(\underset{\sim}{A}\cdot\underset{\sim}{E}_{n-1})\cup(\underset{\sim}{A}\cdot\underset{\sim}{E}_n)]+\cdots \\
&\quad +(-1)^{n+1}\tilde{P}(\underset{\sim}{A}\cdot\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\cdots\cdot\underset{\sim}{E}_n)
\end{aligned}
\tag{2.8.15}
$$

其中 $\underset{\sim}{A}\cdot\underset{\sim}{E}_i\cup\underset{\sim}{A}\cdot\underset{\sim}{E}_j=\underset{\sim}{A}\cdot\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j$, $i,j=1,2\cdots,n$, $i<j$

…

将式(14)代入(15)，即得

$$
\begin{aligned}
\tilde{P}(\underset{\sim}{A}) &= \sum_{i=1}^{n}\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_i)\tilde{P}(\underset{\sim}{E}_i) \\
&\quad -\sum_{\substack{i,j=1\\ i<j}}^{n}\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j)\tilde{P}(\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j) \\
&\quad +\sum_{\substack{i,j,k=1\\ i<j<k}}^{n}\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j\cdot\underset{\sim}{E}_k)\tilde{P}(\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j\cdot\underset{\sim}{E}_k) \\
&\quad -\cdots \\
&\quad +(-1)^{n+1}\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\cdots\cdot\underset{\sim}{E}_n) \\
&\quad \times\tilde{P}(\underset{\sim}{E}_1\cdots\underset{\sim}{E}_n)
\end{aligned}
\tag{2.8.16}
$$

且有

$$
\tilde{P}(\underset{\sim}{E}_i|\underset{\sim}{A})=\frac{\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_i)\cdot\tilde{P}(\underset{\sim}{E}_i)}{\tilde{P}(\underset{\sim}{A})}
\tag{2.8.17}
$$

我们称式(2.8.16)为 F 事件的全概率公式，而式(2.8.17)称为 Bayes 公式。

注 对于式(2.8.12)中考虑的"·"运算，若换为"∧"是不行的。因为对于 $i\neq j$, $i,j=1,2,\cdots,n$,关于 $\underset{\sim}{E}_i\cap\underset{\sim}{E}_j=\phi$ 不再成立。从而使公式(2.8.16)中右端远比公式(2.8.16)要复杂。

例 2.8.1 设

$$E=\{x_1,x_2,x_3,x_4,x_5\}$$

$$\underset{\sim}{E}_1=0.3/x_1+1/x_2+0.2/x_3+0/x_4+1/x_5,$$

$$\underset{\sim}{E}_2=0.6/x_1+0.4/x_2+1/x_3+0.1/x_4+1/x_5,$$

$$\underset{\sim}{E}_3=1/x_1+0.5/x_2+0.9/x_3+1/x_4+0/x_5.$$

显然有

$$(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)=0.18/x_1+0.40/x_2+0.20/x_3+0/x_4+1/x_5,$$

$$(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_3)=0.30/x_1+0.50/x_2+0.18/x_3+0/x_4+0/x_5,$$

$$(\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)=0.60/x_1+0.20/x_2+0,90/x_3+0.10/x_4+0/x_5,$$

$$(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)=0.18/x_1+0.20/x_2+0.18/x_3+0/x_4+0/x_5.$$

又设

$P(x_1)=0.1$, $P(x_2)=0.5$, $P(x_3)=0.1$, $P(x_4)=0.1$, $P(x_5)=0.2$.

于是有

$$\tilde{P}(\underset{\sim}{E}_1)=(0.1)\times(0.3)+(0.5)\times(1)+(0.1)\times(0.2)+(0.1)\times(0)+(0.2\times(1)=0.75,$$

$$\tilde{P}(\underset{\sim}{E}_2)=(0.1)\times(0.6)+(0.5)\times(0.4)+(0.1)\times(1)+(0.1)\times(0.1)+(0.2)\times(1)=0.57,$$

$$\tilde{P}(\underset{\sim}{E}_3)=(0.1)\times(1)+(0.5)\times(0.5)+(0.1)\times(0.9)$$
$$+(0.1)\times(1)+(0.2)\times(0)=0.54,$$

$\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)=0.438$; $\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_3)=0.298$;

$\tilde{P}(\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)=0.260$; $\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)=0.136$.

又设 $\underset{\sim}{A}\subset E$，且

$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1)=0.20$; $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)=0.15$;

$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_2)=0.20$; $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)=0.18$;

$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_3)=0.50$; $\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)=0.15$;

$$\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)=0.10.$$

那么由全概率公式可得

$$\begin{aligned}\tilde{P}(\underset{\sim}{A})=&\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1)\times\tilde{P}(\underset{\sim}{E}_1)+\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_2)\times\tilde{P}(\underset{\sim}{E}_2)\\&+\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_3)\times\tilde{P}(\underset{\sim}{E}_3)\\&-\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)\times\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)\\&-\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_3)\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_3)\\&+\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)\\=&(0.20)\times(0.75)+(0.20)\times(0.57)\\&+(0.50)\times(0.54)-(0.15)\times(0.438)\\&-(0.18)\times(0.298)-(0.15)\times(0.26)\\&+(0.10)\times(0.136)=0.38926.\end{aligned}$$

并且由它可以得到

$$\tilde{P}(\underset{\sim}{E}_1|\underset{\sim}{A})=\frac{\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1)\tilde{P}(\underset{\sim}{E}_1)}{\tilde{P}(\underset{\sim}{A})}=\frac{(0.20)(0.75)}{0.38926}$$
$$=0.38534,$$

同样有

$\tilde{P}(\underset{\sim}{E}_2|\underset{\sim}{A})=0.29286$; $\tilde{P}(\underset{\sim}{E}_3|\underset{\sim}{A})=0.69362$.

这里需提醒读者注意，三个条件概率之和

$$\tilde{P}(\underset{\sim}{E}_1|\underset{\sim}{A})+\tilde{P}(\underset{\sim}{E}_2|\underset{\sim}{A})+\tilde{P}(\underset{\sim}{E}_3|\underset{\sim}{A})\neq 1.$$

这一点与通常概率不一样，原因在于这里有

$$(\underset{\sim}{E}_i\cdot\underset{\sim}{E}_j)\neq\phi,\quad i,j=1,2,3,\quad i\neq j$$

$$(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3)\neq\phi.$$

从而有

$$\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2|\underset{\sim}{A})=\frac{\tilde{P}(\underset{\sim}{A}|\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2)}{\tilde{P}(\underset{\sim}{A})}$$

$$=\frac{(0.15)\cdot(0.438)}{0.38926}=0.16878,$$

$$\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_3|\underset{\sim}{A})=0.13779;$$

$$\tilde{P}(\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3|\underset{\sim}{A})=0.10019;$$

$$\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3|\underset{\sim}{A})=0.03493.$$

由上可得：

$$\begin{aligned}&\tilde{P}(\underset{\sim}{E}_1|\underset{\sim}{A})+\tilde{P}(\underset{\sim}{E}_2|\underset{\sim}{A})+\tilde{P}(\underset{\sim}{E}_3|\underset{\sim}{A})\\&\quad-\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2|\underset{\sim}{A})-\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_3|\underset{\sim}{A})\\&\quad-\tilde{P}(\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3|\underset{\sim}{A})+\tilde{P}(\underset{\sim}{E}_1\cdot\underset{\sim}{E}_2\cdot\underset{\sim}{E}_3|\underset{\sim}{A})\\&=1.\end{aligned}$$

第 三 章

F概率与F测度

本章，着重介绍不分明测度（F 测度）与不分明积分（F积分），以及它们与F概率之间的联系。F测度是作为用不分明的测度计量对象而加以定义的，而F积分是作为用F测度计量不分明对象的手段定义的。

1972年Sugeno首先提出F测度与F积分的概念，后来得到很大的发展。F测度的特征在于用单调性代替了测度定义中的可加性。在此意义下，F测度可看作为概率测度的扩张。我们在这简短的一章里，着重介绍为了以后开展F概率的研究所需要的有关知识。

§1 F 测 度

设Ω为普通集，$\mathscr{P}(\Omega)$为Ω的幂集。

定义 3.1.1 称$\sigma\subset\mathscr{P}(\Omega)$为$\sigma$代数，若

（1）$\Omega\in\sigma$；

（2）$A\in\sigma\Rightarrow A^c\in\sigma$；

（3）$A_n\in\sigma$，$n=1$，2，$\cdots\Rightarrow\bigcup_n A_n\in\sigma$。

(Ω,σ)称为可测空间。

定义 3.1.2 设M为不空集族，若对M中元素的每一个单调序列$\{A_n\}$（递增或递降）恒有

$$\lim_{n\to\infty}A_n\in M,$$

则称M是单调族。

定义 3.1.3 映射 w：$\sigma\to[0,1]$称为不分明测度（简称F测度），若

（1）$w(\phi)=0$，$w(\Omega)=1$；（有界非负性）

（2）$\forall A\subset B\Rightarrow w(A)\leqslant w(B)$；（单调性）

（3）若 $\forall_i$，$A_i\in\sigma$，$i=1$，2，…且单调，则

$$\lim_{i\to\infty} w(A_i)=w(\lim_{i\to\infty} A_i). \qquad \text{（连续性）}$$

(Ω,σ,w)称为 F 测度空间。

F 测度可有多种解释，这里仅举一种：$w(A)$可认为是 Ω 的元素 $x\in A$ 的程度。例如，$w(\phi)=0$ 意味着 $x\in\phi$ 是不可能的，$w(\Omega)=1$ 则表示 $x\in A$ 总是成立的。

例 3.1.1 考虑目测 $\Omega=[0,\ c]$的适当区间的距离．设此不分明的长度的测度为 w，给定 $w(\Omega)=1$，而直观的决定 $A\subset\Omega$ 的 $w(A)$。现在，作为单调集序列可设$A_n=\left[0,c-\frac{c}{n}\right]$，显然 $A_n\subset A_{n+1}$。所以，可给出 A_n和 A_{n+1}的距离使 $w(A_n)\leqslant w(A_{n+1})$。于是这个直观的测度 w 满足 F 测度的条件.

例 3.1.2 概率空间(Ω,σ,P)是 F 测度空间。

例 3.1.3 固定 $x_0\in\Omega$，$\forall A\subset\Omega$，$w(A)=I_A(x_0)$ 是一F测度．

定理 3.1.1 设 w 为$(\Omega,\mathscr{B})$上的 F 测度，其中$\mathscr{B}$为 Ω上的一切 σ Borel 集合系。E，$E'\in\mathscr{B}$，则有

（1）$w(E\cup E')\geqslant w(E)\vee w(E')$ (3.1.1)

（2）$w(E\cap E')\leqslant w(E)\wedge w(E')$ (3.1.2)

证 由 $E\cup E'\supset E$，$E\cap E'\subset E$ 和 w 的单调性结合定义3.1.3，结论是显然的。

为了研究 F 测度的结构及 F 积分的计算．引入 g_λ 测

度。

定义 3.1.4 若 g_λ: $\sigma\to[0,1]$满足条件：

(1) $g_\lambda(\Omega)=1$；

(2) $g_\lambda(A\cup B)=g_\lambda(A)+g_\lambda(B)+\lambda g_\lambda(A)g_\lambda(B)$,

$(A\cap B=\phi)$；

(3) A_n, $n=1, 2, \cdots$单调，$\Rightarrow \lim\limits_{n\to\infty} g_\lambda(A_n)$ $=g_\lambda(\lim\limits_{n\to\infty} A_n)$，称为 g_λ测度。

定理 3.1.2 当 $\lambda>-1$ 时，g_λ 测度是 F 测度。

证 由于 $\Omega\cap\phi=\phi$，则

$$g_\lambda(\Omega)=g_\lambda(\Omega)+g_\lambda(\phi)+\lambda g_\lambda(\Omega)g_\lambda(\phi)$$

但 $g_\lambda(\Omega)=1$，所以 $g_\lambda(\phi)(1+\lambda)=0$，又因 $\lambda>-1$，故 $g_\lambda(\phi)=0$。若 $A\subset B$，则

$$g_\lambda(B)=g_\lambda(A)+g_\lambda(B-A)+\lambda g_\lambda(A)g_\lambda(B-A)$$
$$=g_\lambda(A)+g_\lambda(B-A)(1+\lambda g_\lambda(A))。$$

由于 $\lambda>-1$，故 $g_\lambda(B)\geqslant g_\lambda(A)$，得证。

显然，当 $\lambda=0$ 时，g_λ 测度为概率测度。

从以上可见概率测度可扩展为 g_λ 测度，而 g_λ 测度又是 F 测度。诸如此类的 F 测度很多，如信任测度、可能性测度、似然测度等等都是 F 测度。这里从略（可见 G. Banon 或张文修的书）。

§2 F 积分

运用 Sugeno 的 F 测度可以定义 F 积分。它非常类似于 Lebesgue 积分。设(X,σ)是可测空间，$\mathscr{B}$ 表示$[0, 1]$上的 Borel 集类。h: $X\to[0,1]$称为 σ 可测的，如果对于任意的

$B\in\mathscr{B}$，有

$$h^{-1}(B)=\{x|h(x)\in B\}\in\sigma. \qquad (3.2.1)$$

定义 3.2.1 设 w：$\sigma\to[0,1]$是 F 测度，$A\in\sigma$，h 是 σ 可测的．h 在 A 上的 F 积分为

$$\int_A h(x)\circ w(\cdot)=\sup_{F\in\sigma}[\inf_{x\in F}h(x)\vee w(A\cap F)]$$

定理 3.2.1 F 积分可表示为

$$\int_A h(x)\circ w(\cdot)=\sup_{\alpha\in[0,1]}\min[\alpha,w(A\cap H_\alpha)] \quad (3.2.2)$$

这里 $H_\alpha=\{x|h(x)\geqslant\alpha\}$．

证 令$\alpha_F=\inf\limits_{x\in F}h(x)$，则 $F\subset\{x|h(x)\geqslant\alpha_F\}=H_{\alpha_F}$．于是

$$\sup_{F\in\sigma}[\inf_{x\in F}h(x)\wedge w(A\cap F)]\leqslant\sup_{F\in\sigma}[\alpha_F\wedge w(A\cap H_{\alpha_F})]$$

$$\leqslant\sup_{\alpha\in[0,1]}[\alpha\vee w(A\cap H_\alpha)],$$

又因为$\alpha\leqslant\inf\limits_{\alpha\in H_\alpha}h(x)$，则

$$\sup_{\alpha\in[0,1]}[\alpha\wedge w(A\cap H_\alpha)]\leqslant\sup_{\alpha\in[0,1]}[\inf_{x\in H_\alpha}h(x)\wedge w(A\cap H_\alpha)]$$

$$\leqslant\sup_{F\in\sigma}[\inf_{x\in F}h(x)\wedge w(A\cap F)].$$

得证．

从这个定理可以看出作为 w的定义域至多能考虑为包含单调集列$\{A\cap H_\alpha|\alpha\in[0,1]\}$的集合系．因此，$F$ 积分的定义采用定理 3.2.1 的形式．

类比 Lebesgue 积分很容易得到如下表示：

设$(A_1, A_2, \cdots, A_n)$是 X 的一个划分，即 $A_i\cap A_j=\phi$ $(i\neq j)$，$\bigcup\limits_{i=1}^{n}A_i=X$．又设 S 为简单函数，即

$$S(x)=\sum_{i=1}^{n}\alpha_i I_{A_i}(x) \tag{3.2.3}$$

其中 $\alpha_i\in[0,1]$，I_{A_i}是 A_i的特征函数。记

$$Q_A(S)=\bigvee_{i=1}^{n}[\alpha_i\wedge w(A\cap A_i)] \tag{3.2.4}$$

若 h 是 σ 可测的，记

$$\int_A h(x)dw=\sup_{S<h}Q_A(S). \tag{3.2.5}$$

那么下面的表达式成立：

$$\int_A h(x)\circ w(\cdot)=\int_A h(x)dw. \tag{3.2.6}$$

并且 F 积分具有如下性质：

(1) $0\leqslant\int_A h(x)dw\leqslant 1$;

(2) $h_1\leqslant h_2\Rightarrow\int_A h_1(x)dw\leqslant\int_A h_2(x)dw$;

(3) $w(x)=0\Rightarrow\int_A h(x)dw=0$;

(4) $A\subset B\Rightarrow\int_A h(x)dw\leqslant\int_B h(x)dw$;

(5) $\int_A cdw=c\wedge w(A)\quad(c\in[0,1])$;

(6) $\int_A h(x)dw=\int_X[I_A(x)\wedge h(x)]dw$;

(7) $\int_A(h_1\vee h_2)(x)dw\geqslant\int_A h_1(x)dw\vee\int_A h_2(x)dw$;

(8) $\int_A(h_1\wedge h_2)(x)dw\leqslant\int_A h_1(x)dw\wedge\int_A h_2(x)dw$;

(9) $\int_{A\cup B}h(x)dw\geqslant\int_A h(x)dw\vee\int_B h(x)dw$;

(10) $\int_{A\cap B} h(x)dw \leqslant \int_A h(x)dw \wedge \int_B h(x)dw$;

(11) $\int_X (a\wedge h(x))dw = a\wedge\int_X h(x)dw$;

(12) 对σ可测函数的单调集列$\{h_n\}$，若$\lim\limits_{n\to\infty} h_n = h$，且$h$亦为$\sigma$可测函数，则

$$\int h(x)dw = \lim_{n\to\infty}\int h_n(x)dw.$$

性质(1)至(11)可直接由定义验证.而(12)亦是成立的。事实上，设$\{h_n\}$递增，则$\int h_n(x)dw$递增，且有上界1，于是存在a使$a = \lim\limits_{n\to\infty}\int h_n(x)dw$，又$\int h_n(x)dw \leqslant \int h(x)dw$。从而 $a \leqslant \int h(x)dw$。设$S \leqslant h$，且$S(x) = \sum\limits_{i=1}^{n} a_i I_{A_i}(x)$。对任意$c<1$，令$E_n = \{x | \{h_n(x) \geqslant c\cdot S(x)\}$，显然$\lim\limits_{n\to\infty} E_n = X$且

$$Q_{E_n}(c\cdot S) \leqslant \int_{E_n}(c\cdot S(x))dw \leqslant \int_{E_n} h_n(x)dw$$

$$\leqslant \int h_n(x)dw \leqslant a.$$

又因为$Q(c\cdot S) = \lim\limits_{n\to\infty} Q_{E_n}(c\cdot S)$，$Q(S) = \lim\limits_{C\uparrow 1} Q(c\cdot S)$则$Q(S) \leqslant a$，于是

$$\int h(x)dw = \sup_{S\leqslant h} Q(S) \leqslant a,$$

则证得

$$\lim_{n\to\infty}\int h_n(x)dw = \int h(x)dw.$$

当h_n递降时的情形类似可证。

(13) 若$\int h(x)dw=0 \Rightarrow w\{x \mid h(x)>0\}=0$。

(14) $\int_A h(x)dw=M \Longleftrightarrow w(A\cap H_M)\geqslant M\geqslant w(A\cap H_{\underset{\cdot}{M}})$,

其中$H_M=\{x \mid h(x)\geqslant\alpha\}$，$H_{\underset{\cdot}{M}}=\{x \mid h(x)>\alpha\}$。

(15) 令$H_{\underset{\cdot}{\alpha}}=\{x \mid h(x)>\alpha\}$则

$$\int_A h(x)dw=\sup_{\alpha\in[0,1]}[\alpha\wedge w(A\cap H_{\underset{\cdot}{\alpha}})]$$

(16) 若h可测，则

$$\int_A h(x)dw=\int_0^1 w(H_\alpha\cap A)d\alpha$$

其中右边的测度α为[0,1]上的Lebesgue测度。

(17) 若(X,σ,P)为概率空间，则

$$\left|\int h(x)dP-(L)\int h(x)dP\right|\leqslant\frac{1}{4}.$$

上述(13)—(17)的证明在此从略。

若$\underset{\sim}{A}$是X上的F集合，函数$h(x)$关于w的积分$\int_{\underset{\sim}{A}} h(x)\circ w(\cdot)$定义为

$$\int_{\underset{\sim}{A}} h(x)\circ w(\cdot)=\int_X(\underset{\sim}{A}(x)\wedge h(x))\circ w(\cdot) \quad (3.2.7)$$

显然，对任意$\underset{\sim}{A}$，$\underset{\sim}{B}$，有

$$\int_{\underset{\sim}{A}\cap\underset{\sim}{B}} h(x)\circ w(\cdot)$$

$$\leqslant\int_{\underset{\sim}{A}} h(x)\circ w(\cdot)\wedge\int_{\underset{\sim}{B}} h(x)\circ w(\cdot),$$

$$\int_{\underset{\sim}{A}\cup\underset{\sim}{B}} h(x)\circ w(\cdot)$$
$$\geqslant\int_{\underset{\sim}{A}} h(x)\circ w(\cdot)\vee\int_{\underset{\sim}{B}} h(x)\circ w(\cdot).$$

当 $\underset{\sim}{A}$ 是非 F 集时，这里定义即为定义 3.2.1．F 集合 $\underset{\sim}{A}$ 的 F 测度是

$$w(\underset{\sim}{A})=\int_X \underset{\sim}{A}(x)\circ w(\cdot). \tag{3.2.8}$$

§3　F 积分的计算

为了实际地计算 F 积分的值，有必要就$\mathscr{B}$的一切元素给出与 F 测度的单调性无矛盾的 w 的值。例如，设 X 为 n 个元素构成的有限集合。则$\mathscr{B}$就是 X的一切子集所成的集族$\mathscr{P}(x)$。这时 $\mathscr{P}(x)$中相互不同的单调集序列有 $n!$ 个，就它们独立地定出与单调性矛盾的 w 的值几乎是不可能的．因此，考虑就其中一个单调集序列定出 w 值，而对其它一切元素作出决定 w 值的生成规则。为此目的，可首先考虑 A, $B\in\mathscr{B}$, $A\cap B=\phi$ 时，给出 $w(A)$和 $w(B)$后，由 $w(A)$和$w(B)$生成 $w(A\cup B)$的问题。这时，问题在于生成规则必须满足单调性。为此考虑当 $A\cap B=\phi$ 时，

$$w_\lambda(A\cup B)=w_\lambda(A)+w_\lambda(B)+\lambda w_\lambda(A)w_\lambda(B),$$
$$-1<\lambda<\infty \tag{3.3.1}$$

的测度 w_λ。由于当 $E\subset F$, E, $F\in\mathscr{B}$时，在(3.3.1)中令 $E=A$, $F-E=B$，则有

$$w_\lambda(A\cup B)-w_\lambda(A)=w_\lambda(B)(1+\lambda w_\lambda(A))\geqslant 0 \tag{3.3.2}$$

这恰恰就是由定义3.2.2定义的 g_λ 测度。特别在 $\lambda=0$时，g_λ测度是概率测度。从而可以得出在 R^1 上构造 g_λ 测度。

定义 3.3.1 定义在一维实数空间 R^1上，具有如下性质的函数叫做 F 分布函数：

(1) $0\leqslant H(x)\leqslant 1,\ x\in R^1$；

(2) $x\leqslant y\Rightarrow H(x)\leqslant H(y)$；

(3) $\lim\limits_{x\to a+0} H(x)=H(a)$；

(4) $\lim\limits_{x\to-\infty} H(x)=0,\ \lim\limits_{x\to\infty} F(x)=1$。

定义 3.3.2 设区间$(a,b]\in R^1$，$(a,b]$的测度为

$$g_\lambda((a,b])=\frac{H(b)-H(a)}{1+\lambda H(a)},\quad -1<\lambda<\infty.$$

定理 3.3.1 上述 $g_\lambda((a,b])$是一个 g_λ 测度.

证 定义3.1.4之(1)(3)显然满足，为证得(2)，设 $a,b,c\in R^1$。$a<b<c$，则有

$$\frac{H(c)-H(a)}{1+\lambda H(a)}=\frac{H(b)-H(a)}{1+\lambda H(a)}+\frac{H(c)-H(b)}{1+\lambda H(b)}$$

$$+\lambda\frac{H(b)-H(a)}{1+\lambda H(a)}\cdot\frac{H(c)-H(b)}{1+\lambda H(a)}.$$

从而证得(2)成立。

可见从 H 出发对 R^1 上的一切 Borel 集系$\mathscr{B}$定出 g_λ 的值。因此，给定 $H(x)$就意味着对单调集列$\{(-\infty,x]\}$ 给出了 $g_\lambda((-\infty,x])=H(x)$，结果就由对一个单调集列定出 g_λ 的值而决定了$\mathscr{B}$的一切元素的 g_λ 值。

现在叙述有限集合上 F 积分的计算方法。

设 $X=\{x_1,x_2,\cdots,x_n\}$，$h(x_1)\geqslant h(x_2)\geqslant\cdots\geqslant h(x_n)$，于是 F积分

$$\int_X h(x_i)\circ w(\cdot)=\bigvee_{i=1}^{n}[h(x_i)\wedge w(X_i)].$$

其中 $X_i=\{x_1, x_2, \cdots, x_n\}$，这里，$w$ 的定义域应为 $\mathscr{F}(X)$。所以要求积分的值，必须就 X 的任意子集给出w 的值。但是，这件事一般说来是困难的，所以我们用构成g_λ测度的办法来处理它。

首先设 F 分布函数为

$$H(x_1)\leqslant H(x_2)\leqslant\cdots\leqslant H(x_n)=1$$

定出

$$g_\lambda(X_i)=H(x_i)$$

且定义

$$g_1=H(x_1)$$

$$g_i=\frac{H(x_i)-H(x_{i-1})}{1+\lambda H(x_{i-1})},\qquad 2\leqslant i\leqslant n.$$

对于不相交的子集族 A_i，迭代使用定义 3.1.4 中之(2)，可得

$$g_\lambda\left(\bigcup_i A_i\right)=\begin{cases}\sum\limits_{i=1}^{\infty}g_\lambda(A_i), & \lambda=0\\ \dfrac{1}{\lambda}\left[\prod\limits_{i=1}^{\infty}(1+\lambda g_\lambda(A_i))-1\right], & \lambda\neq 0\end{cases}$$

这里，当 $X=\{x_1,\cdots,x_n\}$时，从 $g_\lambda=g_\lambda(\{x_i\})\in[0,1]$ 使用上面公式，对 X 的任意子集 X' 有

$$g_\lambda(X')=\frac{1}{\lambda}\left[\prod_{x_i\in X'}(1+\lambda g_i)-1\right].$$

因此，从决定单调集列 $X_1\subset X_2\subset\cdots\subset X_n=X$ 的 g_λ 值

的 $H(x_i)$，就可以按照上式定出 X 的任意子集 X' 的 g_λ值。

下面，说明已给定 $h_i = h(x_i)$，$1 \leqslant i \leqslant n$ 时，计算其 F 积分的步骤：

（1）按大小把 h_i排起来，如 $h_{r_1} \geqslant h_{r_2} \geqslant \cdots \geqslant h_{r_n}$，

（2）计算 $H(r_i) = g_{r_i} + H(r_{i-1}) + \lambda g_{r_i} H(r_{i-1})$，

$$H(r_1) = g_1。$$

在此，容易确立

$$H(r_i) = \frac{1}{\lambda}\left[\prod_{j=1}^{i}(1+\lambda g_j) - 1\right]$$

若设 $X_{r_i} = \{x_{r_1}, x_{r_2}, \cdots, x_{r_i}\}$，则 $H(r_i) = g_\lambda(X_i)$，因此，$\int h(x_i) \circ w(\cdot) = \bigvee_{i=1}^{n} [h_{r_i} \wedge H(r_i)]$。

（3）求 $h_{r_i} \wedge H(r_i)$，$1 \leqslant r \leqslant n$ 的最大值．从而得 F 积分值。

例 3.3.1 （F 积分用于多维评价）设要对某种商品进行评价，作为商品评价的要素考虑为 x_1 = “对企业商标的印象”，x_2 = “性能和机能”，x_3 = “设计”，x_4 = “使用方便”这四个项．评价人对这四个项目以何种程度的重视用$g_\lambda(x_i)$表示．而相应于评价的目，对该产品的满意程度用 $h(x_i)$ 表示。为简单计，设 $\lambda = 0$，且 g_λ 和 h 值如表 3.1 所给出．试对该商品作出评价。

这里我们要用 F 积分作评价，设对系统的不分明评价为 g，

$$g = \int h(x) \circ g_\lambda(\cdot)$$

其中 $h(x_i)$表示系统中各元素（项目）x_i的“满意”度的隶属函

数，F 测度 g_λ 表示对各元素 x_i 的重视度。而 g 可理解为：对重视度高的项目，若满意度高的，g 的值也高，反之，g 的值低。

表 3.1

评价 \ 项目	x_1	x_2	x_3	x_4
$h(x_i)$	0.9	0.7	0.4	0.1
g_i	0.5	0.3	0.1	0.1
g'_i	0.2	0.1	0.3	0.4

现在来计算 g 值：(1) $h(x_1)\geqslant h(x_2)\geqslant h(x_3)\geqslant h(x_4)$。(2) $H(x_1)=g_1$ 即 $g_1=0.5$，而 $H(x_2)=g_2+H(x_1)+\lambda g_2H(x_1)=0.3+0.5+0=0.8$；$H(x_3)=0.1+0.8=0.9$；$H(x_4)=0.1+0.9=1$。(3) $h(x_i)\wedge H(i)$，$i=1, 2, 3, 4$ 即 0.5，0.7，0.4，0.1，其中最大者为 0.7，故 $g=0.7$。

当重视度为 g'_i 的场合(如表 3.1 所示)，可得 $g=0.4$。从而说明重视度与满意度之间有密切关系。

§4 F 概率测度

用公理化的方法定义 F 概率，即 F 概率测度，它可看成是 Zadeh 定义的概率的扩展。在 1979 年 E.P. Klement 从任意的可测空间出发，给出一个公理体系。

设 X 是一个任意集合，$I=[0,1]$ 是单位区间，$\mathscr{B}$ 是 I 的 Borel 子集组成的 σ 代数，除非另有说明，总认为 I 具有 $\mathscr{B}$。

定义 3.4.1 $(X,\mathscr{Q})$ 称为 Klement 意义下的 F 可测空间，其意为 $\mathscr{Q}$ 是一个 X 上的 F 集族，满足以下性质：

(1) $\forall \alpha$常数，$\alpha\in \underset{\sim}{\sigma}$；

(2) $\forall \mu\in \sigma$，$\Rightarrow 1-\mu\in \underset{\sim}{\sigma}$；

(3) $\forall (\mu_n)_{n\in N}\in \underset{\sim}{\sigma}^N \Rightarrow \sup\limits_{n\in N}\mu_n\in \underset{\sim}{\sigma}$。

$\underset{\sim}{\sigma}$ 称为 X 上的 $F\sigma$- 代数，其元素称为 F 可测集合（在公理系的那一节，称为 F 事件）。这里的 $N=\{1, 2, \cdots\}$，而 μ 为隶属映射 μ：$X\to[0,1]$。

像通常 σ 代数那样，$F\sigma$ 代数总存在一个最小的 X 上的经典 σ 代数，它由一切 σ 可测函数构成。这个 σ 代数记为 $\boldsymbol{K}(\sigma)$。

显然，若 $\mathscr{A}$ 是 X 上的一个经典的 σ 代数，那么，一切关于$\mathscr{A}$和$\mathscr{B}$可测的函数 μ：$X\to I$的族，是一个 $F\sigma$ 代数，记为 $\zeta(\mathscr{A})$，称为由$\mathscr{A}$生成的 $F\sigma$ 代数。

定义 3.4.2 设$(X,\underset{\sim}{\sigma})$是一个 F 可测空间，F 概率测度是一个映射

$$m: \underset{\sim}{\sigma}\to I$$

满足如下性质：

(4) $m(0)=0$，$m(1)=1$；

(5) $\forall$，μ，$\nu\in \underset{\sim}{\sigma}\Rightarrow m(\mu\vee\nu)+m(\mu\wedge\nu)=m(\mu)+m(\nu)$；

(6) $\forall (\mu_n)_{n\in N}\in \underset{\sim}{\sigma}^N$，$\mu\in \underset{\sim}{\sigma}$

$$(\mu_n)_{n\in N}\uparrow \mu\Rightarrow (m(\mu_n))_{n\in N}\uparrow m(\mu).$$

称$(X,\underset{\sim}{\sigma},m)$是一个 F 概率空间。

由连续性条件(6)不难推出 m 的单调性，即

$$\forall \mu,\nu\in \underset{\sim}{\sigma},\ \mu\leqslant\nu\Rightarrow m(\mu)\leqslant m(\nu).$$

并且，正规化条件(4)与可加性条件(5)结合可推导出

$$\forall \mu,\nu\in \underset{\sim}{\sigma},\ \mu\vee\nu=0\Rightarrow m(\mu+\nu)=m(\mu)+m(\nu).$$

当 $A\cap B=\phi$，且 I_A，$I_B\in \underset{\sim}{\sigma}$ 时，上式给出众所周知的加法

公式

$$m(I_{A\cup B}) = m(I_A) + m(I_B)。$$

然而，在 $\zeta(\mathscr{A})$情形下，可由此推出(5)。

值得提醒的是，一般情况下，m 不满足：

$\forall \alpha$ 常数，$m(\alpha) = \alpha$；

$\forall \mu \in \mathcal{C}$，$m(1-\mu) = 1 - m(\mu)$；

$\forall (\mu_n)_{n\in N} \in \mathcal{C}^N$，$\mu \in \mathcal{C}$，

$(\mu_n)_{n\in N} \downarrow \mu \Rightarrow (m(\mu_n))_{n\in N} \downarrow m(\mu)$。

从 F 概率测度空间$(X, \mathcal{C}, m)$定义可知，Zadeh 的概率定义$\widetilde{P}(\underset{\sim}{A}) = \int \mu_{\underset{\sim}{A}}(x) dP(x)$（其中 $\mu_{\underset{\sim}{A}}(x)$为 $\underset{\sim}{A}(x)$），可知，$\widetilde{P}(\underset{\sim}{A})$是一个 F 概率测度。但其逆不真。

例 3.4.1 设 $X = [0,1]$，$\mathcal{C}$ 为$[0,1]$的一切 Borel子集组成的 $F\sigma$ 代数，且 m 由下式确定：

$$m(\underset{\sim}{A}) = \frac{1}{3} \wedge \mu_{\underset{\sim}{A}}(0) + \frac{2}{3} \vee \mu_{\underset{\sim}{A}}(0) + \frac{2}{3} \wedge \left(\frac{1}{3} \vee \mu_{\underset{\sim}{A}}(1)\right) - 1$$

显然它是一个 F 测度，且为 F 概率测度，但不是 Zadeh 的概率。

现在，设有概率测度 P 存在，使得对一切$\underset{\sim}{A} \in \mathcal{C}$ 有

$$m(\underset{\sim}{A}) = \int \mu_{\underset{\sim}{A}}(x) dP(x).$$

当 $\underset{\sim}{A}$ 由 $\mu_{\underset{\sim}{A}}(x) = I_{[0,\frac{1}{2}]}(x)$定义时，有

$$P\left(\left[0, \frac{1}{2}\right]\right) = \int I_{[0,\frac{1}{2}]}(x) dP(x) = m(\underset{\sim}{A})$$

$$= \frac{1}{3} \wedge 1 + \frac{2}{3} \vee 1 + \frac{2}{3} \wedge \left(\frac{1}{3} \vee 0\right) - 1$$

$$= \frac{1}{3}.$$

而对于由 $\underset{\sim}{\mu_B}(x) = \frac{1}{3} I_{[0,\frac{1}{2}]}(x)$ 确定的 F 事件 $\underset{\sim}{B}$ 有

$$\frac{1}{3} P\left(\left[0, \frac{1}{2}\right]\right) = \int \frac{1}{3} I_{[0,\frac{1}{2}]}(x) dP(x) = m(\underset{\sim}{B})$$

$$= \frac{1}{3}.$$

所以 $P\left(\left[0, \frac{1}{2}\right]\right) = 1$。由此可见，Zadeh的概率定义不是一个积分。那么在什么情况下为一个积分呢？Klemet 作了回答。

设对于每个 $A \in \boldsymbol{K}(\sigma)$，令

$$\mathscr{U}(A) = \{(\mu, \alpha) \in \sigma \times I \mid \mu^{-1}((\alpha, 1]) = A\}$$

定义 3.4.3 称$(X, \underset{\sim}{\sigma})$上一个 F 概率测度 m 满足条件($\mathscr{L}$)是指若对每一个 $A \in \boldsymbol{K}(\sigma)$，$\mathscr{U}(A) \neq \phi$，存在一个数 $P(A) \in I$ 使得

$$\mu^{-1}((\alpha, 1]) = A \Rightarrow \lim_{\beta \downarrow \alpha} \frac{m(\mu \wedge \beta) - m(\mu \wedge \alpha)}{\beta - \alpha} = P(A).$$

注意，并不是每个 F 概率测度都具有条件($\mathscr{L}$)。

例 3.4.2 设 $X = (0, 1]$，$\mathscr{B}$ 为常用的 Borel 集的 σ 代数，令 $\underset{\sim}{\sigma} = \zeta(\mathscr{B})$。设 P_0 与 P_1 是经典概率测度：

$$P_0 = (\{0\}) = P_1(\{1\}) = 1.$$

那么，很容易验证

$$m: \underset{\sim}{\sigma} \to [0, 1],$$

$$\mu \to \int \left[\left(\frac{1}{4} \wedge \mu\right) + \left(\frac{3}{4} \vee \mu\right)\right] dP_0 + \int \frac{3}{4} \wedge \left(\frac{1}{4} \vee \mu\right) dP_1 - 1,$$

是一个 F 概率测度。

令 $\mu=I_{[0,\frac{1}{2}]}\in\mathfrak{C}$。则

$$\mu^{-1}\left(\left(\frac{7}{8},1\right]\right)=\mu^{-1}\left(\left(\frac{1}{2},1\right]\right)=\left[0,\frac{1}{2}\right].$$

假定满足条件($\mathscr{L}$)，对于$\beta>\frac{7}{8}$，我们有

$$\frac{m(\mu\wedge\beta)-m\left(\mu\wedge\frac{7}{8}\right)}{\beta-\frac{7}{8}}=\int I_{[0,\frac{1}{2}]}dP_0=1.$$

所以

$$P\left(\left[0,\frac{1}{2}\right]\right)=1.$$

另外，若$\frac{1}{2}<\beta<\frac{3}{4}$，则

$$\frac{m(\mu\wedge\beta)-m\left(\mu\wedge\frac{1}{2}\right)}{\beta-\frac{1}{2}}=\int I_{[0,\frac{1}{2}]}dP_1=0.$$

从而

$$P\left(\left[0,\frac{1}{2}\right]\right)=0.$$

这个矛盾指出 m 不满足条件($\mathscr{L}$)。

定理 3.4.1 设σ是$\zeta(\mathscr{A})$的子 σ 代数，P 是$(X,\mathscr{A})$上的一个概率测度，则 F 概率测度

$$m:\ \sigma\to[0,1],\quad \mu\to\int\mu dP$$

满足条件($\mathscr{L}$)。

证 任取$A\in\mathbf{K}(\sigma)\subset\mathscr{A}$，使$\mathscr{U}(A)\neq\phi$，且$(\mu,\sigma)\in\mathscr{U}(\mathscr{A})$，

有

$$\lim_{\beta\downarrow\alpha}\frac{m(\mu\wedge\beta)-m(\mu\wedge\alpha)}{\beta-\alpha}$$

$$=\lim_{\beta\downarrow\alpha}\frac{1}{\beta-\alpha}\Big[\beta P(\{\mu>\beta\})-\alpha P(\{\mu>\alpha\})$$

$$+\int_{\{\alpha<\mu\leqslant\beta\}}\mu dP\Big]$$

$$=\lim_{\beta\downarrow\alpha}\frac{1}{\beta-\alpha}\Big[(\beta-\alpha)P\{\mu>\alpha\}-\beta P\{\alpha<\mu\leqslant\beta\}$$

$$+\int_{\{\alpha<\mu\leqslant\beta\}}\mu dP\Big]$$

$$=P(A)+\lim_{\beta\downarrow\alpha}\int_{\{\alpha<\mu\leqslant\beta\}}\frac{\mu-\beta}{\beta-\alpha}dP=P(A).$$

从而，m确实满足条件($\mathscr{L}$)。

定理 3.4.2 设$(X,\mathcal{Q},m)$是一个 F 概率空间，$\mathcal{Q}=\zeta(\mathscr{A})$是一个生成的 $F\sigma$ 代数，则以下命题等价：

(1′) m 是一个积分，即在$(X,\mathscr{A})$上存在一个概率测度P，使得对一切 $\mu\in\mathcal{Q}$ 有

$$m(\mu)=\int\mu dP.$$

(2′) m 满足条件($\mathscr{L}$)；

(3′) m 是线性的，即对一切 μ，$\nu\in\mathcal{Q}$，$\alpha,\beta\in R$ 有

$$\alpha\mu+\beta\nu\in\mathcal{Q}\Rightarrow m(\alpha\mu+\beta\nu)=\alpha m(\mu)+\beta m(\nu).$$

(4′) m 是齐次的，即对所有 $\mu\in\mathcal{Q}$，且对一切$\alpha\geqslant0$，$\alpha\mu\in\mathcal{Q}\Rightarrow m(\alpha\mu)=\alpha m(\mu)$。

(5′) m 是可加的，即对一切 μ，$\nu\in\mathcal{Q}$，有

$$\mu+\nu\in\mathcal{Q}\Rightarrow m(\mu+\nu)=m(\mu)+m(\nu).$$

证 由定理 3.4.1 证得(1)⇒(2)。

由(1′)⇒(3′)和(3′)⇒(5′)是明显的。仅需证明(2′)⇒

(4)′，(4′)⇒(1′)及(5′)⇒(4′)。现证由(2′)⇒(4′)；假定 m 满足条件($\mathscr{L}$)，且对于任意 $A\in\mathscr{A}$，考虑函数 f_A: $I\to I$，$\alpha\to m(\alpha\wedge I_A)$。这是在(0,1]上左连续函数，而且在[0，1)上右连续，所以它在[0,1]上连续。

对每个 $\alpha\in[0,1)$，f_A的右导数等于数 $P(A)$，从而 f_A在(0,1)上是可微的。由于对 $\alpha<\beta\in(0,1)$有

$$f_A(\beta)-f_A(\alpha)=\int_\alpha^\beta f'_A(x)dx=(\beta-\alpha)P(A).$$

f_A的连续允许我们考虑 $\alpha\to 0$ 的极限，得

$$\beta P(A)=f_A(\beta),$$

又让 $\beta\to 1$ 得

$$P(A)=f_A(1)=m(I_A).$$

为了证得对 $\mu\in\mathcal{Q}$ 的齐次性，我们首先考虑一个可测的阶梯函数

$$\mu=\sum_{i=1}^n\alpha_i I_A \qquad (A_i\text{ 两两不相交})$$

若 $\alpha\mu\in\mathcal{Q}$，那么用 $m(I_{A\cup B})=m(I_A)+m(I_B)$得

$$m(\alpha\mu)=m\left(\sum_{i=1}^n\alpha\alpha_i I_A\right)=\sum_{i=1}^n m(\alpha\alpha_i I_A)$$

$$=\sum_{i=1}^n\alpha m(\alpha_i I_A)=\alpha m(\mu).$$

现在，关于任一 $\mu\in\mathcal{Q}$，存在一个阶梯函数的递增序列 $(S_n)_{n\in N}\in\mathcal{Q}$，使得

$$\mu=\sup_{n\in N}S_n$$

其次，若 $\alpha\mu\in\mathcal{Q}$ 且使用 F 概率测度的连续性得

$$m(\alpha\mu)=m\ (\sup_{n\in N}\alpha S_n)=\sup_{n\in N}m(\alpha S_n)$$

$$=\alpha\sup_{n\in N} m(S_n)=\alpha m(\mu).$$

所以 m 是齐次的。

现在来证(4′)⇒(1′)。假定 m 是齐次的。那么

$$P: \mathscr{A}\to I,\ A\to m(I_A)$$

是一个概率测度。事实上，由定义中正规性知 $P(\phi)=0$，$P(X)=1$。且由连续性及 $m(I_{A\cup B})=m(I_A)+m(I_B)$ 知 P 是 σ 可加的。对于一个阶梯函数

$$\mu=\sum_{i=1}^{n}\alpha_i I_{A_i}\qquad (A_i\text{ 两两不相交})$$

由齐次性及 $m(I_{A\cup B})=m(I_A)+m(I_B)$ 得

$$m(\mu)=m\left(\sum_{i=1}^{n}\alpha_i I_{A_i}\right)=\sum_{i=1}^{n}\alpha_i P(A_i)=\int\mu dP$$

对于任一 $\mu\in\mathscr{G}$，考虑一个阶梯函数的增序列 $(S_n)_{n\in N}\in\sigma^N$，使 $\mu=\sup_{n\in N} S_n$，那么

$$m(\mu)=m(\sup_{n\in N} S_n)=\sup_{n\in N}\int S_n dP=\int\mu dP.$$

所以 m 是一个积分。

最后证(5′)⇒(4′)：假定 m 是可加的.考虑 $\alpha>0$，$\mu>0$，使得 $\alpha\mu\in\sigma$，若 $\alpha=0$，由 m 的正规性得 $m(0\cdot\mu)=0\cdot m(\mu)$。若 $\alpha>0$，对一切 r，$S\in N$ 有

$$\frac{r}{S}\leqslant\alpha\quad\Rightarrow\quad\frac{r}{S}\mu\in\sigma,$$

且由 m 的可加性有 $m\left(\frac{r}{S}\mu\right)=r\cdot m\left(\frac{1}{S}\mu\right)=\frac{r}{S}m(\mu)$.

现在对一个满足 $\alpha=\sup_{n\in N}\alpha_n$ 的有理数递增序列 $(\alpha_n)_{n\in N}$，得

$$m(\alpha\mu)=m(\sup_{n\in N}\alpha_n\mu)=\sup_{n\in N} m(\alpha_n\mu)=\alpha m(\mu).$$

所以 m 是齐次。

至此定理全部得证。

§5　FP 测度及其扩展

在 1986 年，K.Piasecki 研究了 Klements 及 Smets 等人所定义的 F 事件的概率，认为它们存在不足之处。这个不足表现为仅在特殊情形有 Bayes 公式。为此他引进FP测度。

设 X 是一个脆弱集(Crisp set)，即隶属函数仅取值 0 与 1 之集，或说是一普通集。又设 $\sigma=\{\mu: X\to[0,1]\}$ 为 $F\sigma$ 代数。即σ 包含$\{0,1\}$(亦记作$\{0_X,1_X\}$)，且关于逆及可列并运算封闭。

定义 3.5.1　不含$\left[\frac{1}{2}\right]_X X:\to\left\{\frac{1}{2}\right\}$的 F 代数(σ 代数)称为软 F 代数(σ 代数)。

定义 3.5.2　由

$$\forall\mu\in\mathscr{F}(X).\quad \mathscr{K}(\mu)=\left\{x\,|\,x\in X,\ \mu(x)>\frac{1}{2}\right\}$$

所定义的映射 $\mathscr{K}:\mathscr{F}(X)\to 2^X$，称为一个非空支集；而由

$$\forall\mu\in\mathscr{F}(X)\quad \mathscr{K}^*(\mu)=\left\{x\,|\,x\in X,\ \mu(x)=\frac{1}{2}\right\}$$

所定义的映射 $\mathscr{K}^*:\mathscr{F}(X)\to 2^X$ 称为未定元素(ill-define elements)的支集。2^X为 X 的幂集。

记

$$L(\mu)=\mathscr{K}(\mu)\cup\mathscr{K}^*(\mu),$$

$$\boldsymbol{K}(\Phi)=\{A\,|\,A\in 2^X,\exists\mu\in\Phi:\mathscr{K}\subset A\subset L(\mu)\},$$

$$\boldsymbol{K}^*(\Phi)=\{A\,|\,A\in 2^X,\ \exists\mu\in\Phi: A=\mathscr{K}^*(\mu)\}.$$

其中 $\Phi\subset\mathscr{F}(X)$，下同

$$\boldsymbol{E}(\Phi)=\{\mu\,|\,\mu\in\mathscr{F}(X),\ \exists(A,B)\in\boldsymbol{K}^2(\Phi):$$

$A \subset B$, $\mathscr{K}(\mu)=A$ 且 $L(\mu)=B\}$

$$\boldsymbol{E}^*(\Phi)=\{\mu \mid \mu \in \boldsymbol{E}(\Phi): \exists A \in \boldsymbol{K}^*(\Phi); \ \mathscr{K}^*(\mu) \subset A\}.$$

我们可以证明以下命题成立。

(1) 若 $\sigma \subset \mathscr{F}(X)$是$(X)$的一个 $F\sigma$ 代数，则 $\boldsymbol{K}(\sigma)$ 是一个含 $\boldsymbol{K}^*(\sigma)$的脆弱 σ 代数。

(2) 若 σ 是一个 $F\sigma$ 代数，则 $\boldsymbol{E}(\sigma)$是一个包含σ的 $F\sigma$ 代数。而且 $\boldsymbol{K}(\sigma) \subset \boldsymbol{K}(\boldsymbol{E}(\sigma))$。

(3) 若 σ 是一个软 $F\sigma$ 代数，则 $\boldsymbol{E}(\sigma)$是一个含 σ 的软 $F\sigma$ 代数。

定义 3.5.3 具有以下性质的映射$P: \sigma \to R^+ \cup \{0\}$，称为 FP-测度：

(i) 对任一 $\mu \in \sigma$，$P(\mu \vee (1-\mu))=1$； (P.1)

(ii) 对任一有限或无限序列 $\{\mu_n\} \in \sigma$，若对于使得 $i \neq j$ 的每一对(i,j)有 $\mu_i \leqslant 1-\mu_j$，则

$$P(\sup_n\{\mu_n\})=\sum_i P(\mu_n). \tag{P.2}$$

由下面的定理可以看出，每一个 FP-测度是一个 F 概率测度。

定理 3.5.1 任何 FP-测度是一个满足下面条件的单调函数 P，对任意$(\mu,\nu) \in \sigma^2$有

$$P(\mu \wedge (1-\mu))=0, \tag{3.5.1}$$

$$P(1-\mu)=1-P(\mu) \tag{3.5.2}$$

$$P(\mu \vee \nu)+P(\mu \wedge \nu)=P(\mu)+P(\nu) \tag{3.5.3}$$

证 因为 $(\mu,\nu) \in \sigma^2$，使 $\mu \leqslant 1-\nu$。从而 $\mu \vee (1-\mu)$ 和 $\mu \wedge (1-\mu)$之间有

$$P(\mu \vee (1-\mu))=P[(\mu \vee (1-\mu)) \vee (\mu \wedge (1-\mu))]$$

$$= P(\mu \vee (1-\mu)) + P(\mu \wedge (1-\mu))$$

所以(3.5.1)成立。由$(P.1)$及$(P.2)$有

$$1 = P[\mu \vee (1-\mu)] = P(\mu) + P(1-\mu)$$

所以(3.5.2)成立。又因为 F 集合

$$\nu \wedge \mu \leqslant \mu = 1-(1-\mu) \leqslant 1-\nu \wedge (1-\mu) \qquad (*)$$

$$1-\mu \vee (1-\mu) \leqslant 1-(1-\mu) = \mu \leqslant \mu \vee \nu = 1-(1-\mu \vee \nu).$$

所以有

$$\begin{aligned} P(\mu) + P(\nu \wedge (1-\mu)) &= P(\mu \vee (\nu \wedge (1-\mu))) \\ &= P[(\mu \vee \nu) \wedge (\mu \vee (1-\mu))] \\ &= 1 - P[(1-\mu \vee \nu) \vee (1-\mu \vee (1-\mu))] \\ &= 1 - (P(1-\mu \vee \nu) + P[1-\mu \vee (1-\mu)]) \\ &= P(\mu \vee \nu) + P(\mu \vee (1-\mu)) - 1 = P(\mu \vee \nu). \end{aligned}$$

$$\begin{aligned} P(\nu \wedge (1-\mu)) + P(\mu \wedge \nu) &= P[(\nu \wedge (1-\mu)) \vee (\mu \wedge \nu))] \\ &= P[\nu \wedge (\mu \vee (1-\mu))] \\ &= 1 - P[(1-\nu) \vee (1-\mu \vee (1-\mu))] \\ &= 1 - P[(1-\nu) \vee (\mu \wedge (1-\mu))] \\ &= 1 - (P(1-\nu) + P(\mu \wedge (1-\mu))) = P(\nu). \end{aligned}$$

于是得到(3.5.3)。而由 $\mu \leqslant \nu$，可得$(*)$及

$$\mu \wedge (1-\mu) \leqslant \mu \leqslant 1-(1-\nu).$$

于是由定义 3.5.3(ii)得

$$\begin{aligned} P(\mu) \leqslant P(\mu) + P(\nu \wedge (1-\mu)) &= P[\mu \vee (\nu(1-\mu))] \\ &= P[\nu \wedge (\mu \vee (1-\mu))] \\ &= 1 - P[(1-\nu) \vee (1-\mu \vee (1-\mu))] \\ &= 1 - P[(1-\nu) \vee (\mu \wedge (1-\mu))] \\ &= 1 - [P(1-\nu) + P(\mu \wedge (1-\mu))] = P(\nu). \end{aligned}$$

从而证得定理。

定理 3.5.2 任何一个 FP-测度 P 对每对 $(\mu,\nu)\in\sigma^2$ 使 $\mu\leqslant\nu$ 满足条件

$$P(\nu\wedge(1-\mu))=0 \quad \text{当且仅当 } P(\mu)=P(\nu).$$

证 若 $\mu\leqslant\nu$，则 F 集 $\nu\wedge(1-\mu)$ 有

$$\nu\wedge(1-\mu)\geqslant 1-(\nu\wedge(1-\mu)).$$

结合$(P1)$，(3.5.1)和(3.5.3)有

$$\begin{aligned}0&=P[\mu\wedge(1-\mu)]\\&=P(\nu)+P(1-\mu)-P[\nu\wedge(1-\mu)]\\&=P(\nu)-P(\mu).\end{aligned}$$

从而得证。

定理 3.5.3 任一 FP-测度 P 对每个非减序列 $\{\mu_n\}\uparrow\mu\in\sigma$，有$\{P(\mu_n)\}\uparrow P(\mu)$。

由于这个定理的证明比较长，我们就不证了．详细证明可在 K.Piasecki 的 1985 年的文中找到。

如果对于 FP-测度空间(X,σ,P)使用限制：

(a) X 称为基本事件集；

(b) 从 σ 每一个映射称为一个 F 事件；

(c) 值 $P(\nu)$称为 F 事件 ν 的概率；

(d) 每个概率为正的 F 事件是可能的；

(e) 每个概率为零的 F 事件是不可能的；

(f) 每个概率为 1 的 F 事件是必然的；

(g) F 事件的补称为逆事件．

那么在该空间上开展概率的理论，这个软 F 概率空间的理论非常类似于经典概率论．以下记

映射 $P:\boldsymbol{K}(\sigma)\to[0,1]$，由下式给定

$$\forall A\in\boldsymbol{K}(\sigma),\ P(A)=p(\mu),\ \boldsymbol{K}(\mu)\subset A\subset L(\mu).$$

映射 $\overline{p}:\boldsymbol{E}(\sigma)\to[0,1]$，由下式给定

$$\forall A\in\boldsymbol{K}(\sigma),\ P(A)=p(\mu),\ \boldsymbol{K}(\mu)\subset A\subset\boldsymbol{L}(\mu), \tag{3.5.4}$$

$$\overline{p}(\mu)=P[\mathscr{K}(\mu)] \tag{3.5.5}$$

其中 p 为$p:\sigma\to R^{+}\cup\{0\}$，是 FP-测度。

定理 3.5.4 映射 $\overline{p}$ 对于每个 $\mu\in\boldsymbol{E}(\sigma)$ 是 Klement 意义下的一个 F 概率测度(见定义 3.4.2)。

证 我们有 $\overline{p}(\phi_X)=P(K(\phi_X))=P(\phi)=0$，且 $\overline{p}(1_X)=P(K(1_X))=P(\Omega)=1$。因此，定义 3.4.2 之(4)成立。设 $(\mu,\nu)\in\boldsymbol{E}^2(\sigma)$。那么由(3.5.5)得

$$\begin{aligned}\overline{p}(\mu\vee\nu)&=P(\mathscr{K}(\mu\vee\nu))=P(\mathscr{K}(\mu)\cup\mathscr{K}(\nu))\\&=P(\mathscr{K}(\mu))+P(\mathscr{K}(\nu))-P(\mathscr{K}(\mu)\cap\mathscr{K}(\nu))\\&=\overline{p}(\mu)+\overline{p}(\nu)-P(\mathscr{K}(\mu\wedge\nu))\\&=\overline{p}(\mu)+\overline{p}(\nu)-\overline{p}(\mu\wedge\nu).\end{aligned}$$

从而满足定义 3.4.2 的(5)。进而若 $\{\mu_n\}\subset\boldsymbol{E}(\sigma)$，且 $\{\mu_n\}\uparrow\mu\in\boldsymbol{E}(\sigma)$，则$\{\mathscr{K}(\mu_n)\}\uparrow\mathscr{K}(\mu)$。因此，

$$\{\overline{p}(\mu_n)\}=\{P(K(\mu_n))\}\uparrow P(\mathscr{K}(\mu))=\overline{p}(\mu).$$

Smets 将定义 3.4.2 中(6)换为条件：

$$\forall\{\mu_n\}\in\mathcal{G},\ \{\mu_n\}\downarrow\mu\in\mathcal{G}\Rightarrow\{m(\mu_n)\}\downarrow m(\mu). \tag{3.5.6}$$

定理 3.5.5 映射 $\hat{p}:\boldsymbol{E}(\sigma)\to[0,1]$，由(3.5.4)及

$$\hat{p}(\mu)=P(L(\mu))\qquad \mu\in\boldsymbol{E}(\sigma) \tag{3.5.7}$$

给定。则 $\hat{p}$ 是一个满足(3.5.6)的 F 概率测度。

证 类似于定理 3.5.4 可证。

比较 $\overline{p}$ 和 $\hat{p}$。因为 $\overline{p}\left(\left[\frac{1}{2}\right]_X\right)=0\neq\hat{p}\left(\left[\frac{1}{2}\right]_X\right)$，$\overline{p}$ 和 $\hat{p}$ 是 $\boldsymbol{E}(\sigma)$ 上不同的 F 概率测度。P 的单调性包含 $\overline{p}(\mu)$

$\leqslant \hat{p}(\mu)$。如果考虑 F 半测度 $\widetilde{m}: \mathcal{Q} \to [0,1]$(定义为一个单调映射)。显然任一 F 概率测度是一个 F 概率半测度。

定理 3.5.6 若对每个 $\mu \in \boldsymbol{E}^*(\sigma)$，有 $\widetilde{m}(\mu) = \bar{p}(\mu) = \hat{p}(\mu)$ 的 F 概率半测度，那么它满足 $\bar{p}(\mu) \leqslant \widetilde{m}(\mu) \leqslant \hat{p}(\mu)$，$\mu \in \boldsymbol{E}(\sigma)$。

证 令 $\mu \in \boldsymbol{E}(\sigma)$。设 $\widetilde{m}(\mu) > \hat{p}(\mu)$，按 $\boldsymbol{K}(\Phi)$ 和 $\boldsymbol{E}(\Phi)$ 的含义这里存在 $\nu \in \sigma$ 使 $\mathscr{K}(\nu) \subset L(\mu) \subset L(\nu)$。因此 $\mathscr{K}^*(\mu \vee \nu) \subset \mathscr{K}^*(\nu) \in \mathscr{K}^*(\sigma)$。故 $\mu \vee \nu \in \boldsymbol{E}^*(\sigma)$。结合 $\hat{p}$，得

$$\begin{aligned} P(L(\nu)) = \hat{p}(\nu) = p(\nu) = P(L(\mu)) = \hat{p}(\mu) < \widetilde{m}(\mu) \\ \leqslant \widetilde{m}(\mu \vee \nu) = \hat{p}(\mu \vee \nu) = P(L(\mu \vee \nu)) \\ = P(L(\mu) \vee L(\nu)) = P(L(\nu)). \end{aligned}$$

若设 $\widetilde{m}(\mu) < \bar{p}(\mu)$。同上理由，存在 $\varphi \in \sigma$ 使 $\mathscr{K}(\varphi) \subset \mathscr{K}(\mu) \subset L(\varphi)$。因此 $\mathscr{K}^*(\mu \vee \varphi) \subset \mathscr{K}^*(\varphi) \in \boldsymbol{K}^*(\sigma)$。所以 $\mu \wedge \varphi \in \boldsymbol{E}^*(\sigma)$。从而得

$$\begin{aligned} P(\mathscr{K}(\varphi)) = \bar{p}(\varphi) = p(\varphi) = P(\mathscr{K}(\varphi)) = \bar{p}(\mu) \\ > \widetilde{m}(\mu) \geqslant \widetilde{m}(\mu \wedge \varphi) = \bar{p}(\mu \wedge \varphi) \\ = P(\mathscr{K}(\mu \wedge \varphi)) = P(\mathscr{K}(\mu) \cap \mathscr{K}(\varphi)) \\ = P(\mathscr{K}(\varphi)). \end{aligned}$$

以上皆得出矛盾结论。从而证得 $\bar{p}(\mu) \leqslant \widetilde{m}(\mu) \leqslant \hat{p}(\mu)$。

通过以上诸定理可以看出：FP-测度是 Sugeno 测度、Khalili 的 F 测度和 Klement 的 F 概率测度，而且很容易验证 Zadeh 的 F 事件的概率是一个 FP-测度(仅在脆弱情形)。另外，由通常测度生成 FP-测度的扩展，已在定理 3.5.4–5 中谈及与 F 概率间的关联。这一切都有着广阔的研究前景。

§6 Bayes 公式与FP测度

我们目前来研究在F概率测度基础上的 Bayes 公式.设$\sigma=\{\mu:\Omega\to[0,1]\}$是$F\sigma$-代数，$\tilde{\sigma}$是其软$F\sigma$-代数，$m$为映射$m:\sigma\to[0,1]$，是$F$概率测度。

定义 3.6.1 由等式

$$c(\mu|\nu,m)=\frac{m(\mu\wedge\nu)}{m(\nu)} \tag{3.6.1}$$

所定义的映射$c(\cdot|\nu,m):\sigma\to[0,1]$，其中对任何$F$概率测度$m$和每个$\nu\in\sigma$，使得$m(\nu)\neq0$，称为在给定$\nu$条件下的条件$F$概率。

定义 3.6.2 满足以下条件的Ω上的F子集序列$\{\mu_m\}\in\sigma^N$称为Ω的一个 Bayes 划分：

(R_1) μ_n是两两W-可分的，即对于每对正整数(k,l)，$k\neq1$，有$\mu_k\leqslant1-\mu_l$；

(R_2) $m(\sup\limits_n\{\mu_n\})=1$；

(R_3) 对每个自然数n有$m(\mu_n)>0$。

定理 3.6.1 若$\{\nu_q\}$和$\{\mu_r\}$是由FP-测度p确定的 Bayes F划分。则对任意的$(\mu_l,\nu_k)\in\tilde{\sigma}^2$有

$$c(\nu_k|\mu_l,p)=\frac{p(\nu_k)\cdot c(\mu_l|\nu_k,p)}{\sum\limits_q p(\nu_q)\cdot c(\mu_l|\nu_q,p)}$$

由于任何一个FP-测度是一个F概率测度，而 Bayes F划分对每个FP-测度包含上升的子集存在，因此任何FP-测度满足 Bayes 公式。

由(3.6.1)式可见，单元素序列$\{1_\Omega\}$对于任一F概率测度来说是一个 Bayes F划分。一般，若对于F概率测度m

有 $m(\mu)=1$，那么$\{\mu\}$是由测度 m 生成的一个 Bayes F 划分。众所周知，对于 Bayes 推断方法，单元素 Bayes 划分没有什么用处，我们看到单元素序列 $\{1_\Omega\}$ 是由 Zadeh 的 F 事件的概率测度关于 $F\sigma$-代数 $\sigma_c=\{[c]_\Omega:\Omega\to\{c\};\ c\in[0,1]\}$确定的唯一的 Bayes F 划分。所以一个 F 概率测度生成一个单元素 Bayes F 划分存在。

定义 3.6.3 如果一个 F 概率测度能生成一个双元素或更多元素的 Bayes F 划分，则称它为一个标准 F 概率测度。

定理 3.6.2 一个标准 F 概率测度满足 Bayes 公式当且仅当它是 FP-测度。

证 设 m 是一标准 F 概率测度. 它不满足 Bayse 公式，则存在 Bayes 划分$\{\nu_q\}$、$\{\mu_r\}$和正整数对(K,l)使得

$$c(\nu_k|\mu_l,m)\neq\frac{m(\nu_k)\cdot c(\mu_l|\nu_k,m)}{\sum_q m(\nu_q)\cdot c(\mu_l|\nu_q,m)}.$$

同条件 F 概率的定义一起指出

$$m(\mu_l)\neq\sum_q m(\mu_l\wedge\nu_q).$$

若$\{\nu_q\}$是单元素序列则上式不成立。有 $m(\phi\wedge\eta)=m(\phi)$，对每个$(\phi,\eta)\in\sigma^2$，使 $m(\eta)=1$。所以，由(R_2)得到

$$\sum_q m(\mu_l\wedge\nu_q)\neq m(\mu_l\wedge\sup_q\{\nu_q\})=m(\sup_q\{\mu_l\wedge\nu_q\}).$$

可见 F 概率测度 m 关于满足条件(R_1)的序列 $\{\mu_l\wedge\nu_q\}$ 不能实现条件 P_2(即 FP 测度的第二个条件)。所以，映射 m 不是一个 FP 测度。结合 FP 测度的 Bayes 公式表示，证得定理.

可见 FP 测度是一个对于 Bayes 推断方法有用的 F 概率测度。

第四章

F 随机变量

前两章我们建立了 F 事件的概率空间($\Omega, \mathscr{K}, P$)，从原则上说，已经完全建立了 F 事件概率论的逻辑基础。正如古典概率论一样，还需要在此基础上给出一些基本概念，首先是不分明(F)随机变量的概念。

F 随机变量的概念曾由 Zadeh, Neguyen, Kandel, Byan, Kwakermaak, Stem, Talati等人所探讨. 虽然观点稍有差异，但结构相似。本章将在回顾经典概率中随机变量概念的基础上介绍 F 随机变量及其性质。

§1 随机变量

回顾经典概率论中，随机变量是一般测度论中可测函数的特殊情形。那时我们认为随机变量本身的描述只用到样本空间 Ω 和σ代数 $\mathscr{F}$，因而把随机变量看作定义在可测空间($\Omega, \mathscr{F}$)上的可测函数的特例。

定义 4.1.1 设($\Omega, \mathscr{F}, P$)是一概率空间，$\xi=\xi(\omega)$，$\omega\in\Omega$，是定义在 Ω 上的实值函数，且对任何实数 x

$$\{\omega: \xi(\omega)<x\}\in\mathscr{F}. \tag{4.1.1}$$

则称 ξ 是($\Omega, \mathscr{F}, P$)上的一个实随机变量，若 $\xi=\eta+i\zeta$，而 η，ζ 是($\Omega, \mathscr{F}, P$)上的实随机变量，则称 ξ 为($\Omega, \mathscr{F}, P$)上的一个复随机变量。

该定义显然可以扩展至 n 维随机向量($\xi_1, \xi_2, \cdots, \xi_n$)。

对于一维情形来说，实际上要求集合$\{\omega:\xi(\omega)<x\}$应有概率存在。换句话说，它是$\mathscr{F}$可测的。

设$L=[0,1]$，经典概率空间为$(\Omega,\mathscr{F},P)$。记

$$\mathscr{A}=\{\chi\mid\chi:\Omega\to[0,1],\ \chi^{-1}([\alpha,1])\in\mathscr{F},\ \forall\alpha\in L\}.$$

$$\mathscr{B}=\{f\mid f:L\to\mathscr{F},\ f(0)=\phi,\ f(\bigvee_n\alpha_n)=\bigcup_n f(\alpha_n)\}.$$

其中$\mathscr{A}$为该概率空间上一切随机变量的集合，由于$\mathscr{A}$与$\mathscr{B}$同构，从而$\mathscr{A}$的表现是作为$(A_\alpha)_{\alpha\in N}\subset\mathscr{K}$，$A_0=\phi$，$A_{\bigvee_n\alpha_n}=\bigcup_{n\in N}A_{\alpha_n}$，$\alpha_n\in L$，$n\in N$的集合的全体而给出的。如此，用$f$映射将随机变量表现出来，故可以讨论随机变量的概率问题。并且注意χ在F情形是隶属函数，而$\chi^{-1}([\alpha,1])=\{\omega\mid\chi(\omega)\geq\alpha\}$为$\alpha$水平集，它属于$\mathscr{F}$，即$\mathscr{F}$-可测。

§2 KF集及SF变量

为了研究F随机变量，先介绍Kwakernaak关于F集合的不同看法以及S.Nahmias的F变量。

4.2.1 Kwakernaak的F集。

设X是一个称为基本空间的普通集。函数$\mu:X\to[0,1]$是隶属函数。又设映射$a:X\to M$，M是一个“抽象的全域”，其上定义了运算$\wedge$(并)与$\vee$(或)，是一个完全分配格.对每个$x\in X$，指定了一个命题$a(x)$，所以M是一个命题的集。在M上，我们定义一个函数$t:M\to[0,1]$，对于给定的$a\in M$，数$t(a)$是语句a的真值。而$\mu(x)=t(a(x))$，即隶属函数$\mu(x)$相应于语句$a(x)$的真值。关于t，Bellman和Giertz增加一些限制(见第二章§6)：

(1) 存在函数 f 和 g，将 $[0,1]\times[0,1]$ 映射到 $[0,1]$ 中，使对一切 $a\in M$，$b\in M$，有

$t(a\wedge b)=f(t(a),\ t(b))$，$t(a\vee b)=g(t(a),\ t(b))$，

(2) $f(x,y)$ 和 $g(x,y)$ 连续且关于 x 非降，

(3) $f(x,x)$ 和 $g(x,x)$ 就 x 严格递增，

(4) $f(x,y)\leqslant\min(x,y)$，$g(x,y)\geqslant\max(x,y)$，对一切 $x\in[0.1]$，$y\in[0,1]$。

(5) $f(1,1)=1$，$g(0,0)=0$.

证明了 f 和 g 由

$$f(x,y)=\min(x,y),\quad g(x,y)=\max(x,y)$$

唯一确定。

设 Λ 是一指标集，对于语句 $a_\lambda\in M$，$\lambda\in\Lambda$ 的任一子集有

$$t\Big(\bigwedge_{\lambda\in\Lambda}a_\lambda\Big)=\inf_{\lambda\in\Lambda}t(a_\lambda),\quad t\Big(\bigvee_{\lambda\in\Lambda}a_\lambda\Big)=\sup_{\lambda\in\Lambda}t(a_\lambda).$$

定义 4.2.1 三元总体 $\alpha=(X,\mu,a)$ 称为不分明集 α，简称 α 为 KF 集合。

如此，一个不分明集是语句与每个成员语句的真值一起的指标集。这个观点与前面引进的不分明集的看法稍微有些差别，但不是本质的。故记为 KF 集合，以示其差异。

给定一个 KF 集 $\alpha=(X,\mu,a)$。令 ϕ 是一个函数映射 $X\to Y$，其中 Y 是另一个普通集。考虑下述语句：对任何一个 $y\in Y$，存在一个元素 $x\in X$，使得 $a(x)$ 成立，且 $\phi(x)=y$。该语句可以定义为

$$b(y)\triangleq\bigvee_{x\in X}[a(x)\wedge(\phi(x)=y)].$$

这里包括了 M 中的语句($\phi(x)=y$)。复合语句 $b(x)$ 的真值由真值函数的属性决定：

$$
\begin{aligned}
t(b(y)) &= t\left(\bigvee_{x\in X}[a(x)\wedge(\phi(x)=y)]\right) \\
&= \sup_{x\in X} t(a(x)\wedge(\phi(x)=y)) \\
&= \sup_{x\in X} \min[t(a(x)),\ t(\phi(x)=y)] \\
&= \sup_{x\in X:\phi(x)=y} \mu(x),
\end{aligned}
$$

因为，若 $\phi(x)=y$ 时，$t(\phi(x)=y)=1$，若 $\phi(x)\neq y$ 时，$t(\phi(x)=y)=0$。

从而，我们可以得到

定义 4.2.2 令 $t(b(y))=v(y)$，得不分明集 (Y, v, b)。称它为不分明集 $\alpha=(X, \mu, a)$在映射 ϕ 下的像(Y 中)，记作 $\phi(\alpha)$。

注 这里将上面所涉及的映射系统归纳一下是有益的：

X：普通集，基本空间；

μ：隶属函数，是映射 $\mu:X\to[0,1]$；

M：语句集；

a：命题；

Y：另一普通集，

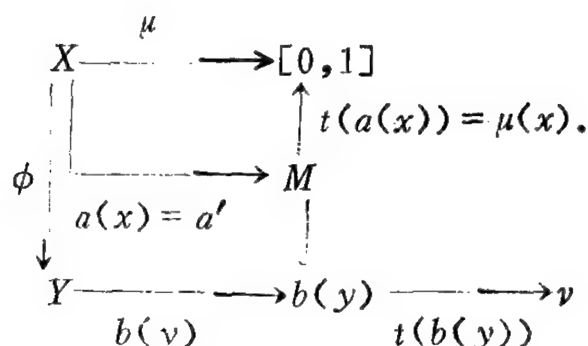
$t(x)$：x 的真值。

于是

$\alpha=(X,\mu,a)$：不分明集，

(Y,v,b)：不分明集。

且他们之间有如下的关系：



从而

$$(X,\mu,a)\xrightarrow[t(b)=\nu]{\phi(a):}(Y,\nu,b).$$

定理 4.2.1 设 $\alpha=(X,\mu,a)$是一个 KF 集，并设 ϕ 和 ψ 分别是函数映射 $\phi: X\to Y$ 和 $\psi: Y\to Z$，则

$$\psi(\phi(\alpha))=(\psi\circ\phi)(\alpha).$$

即 $\phi(\alpha)$在映射 ψ 之下 Z 的像与 α 在复合映射$\psi\circ\phi: X\to Z$之下 Z 中的像相同。这里

$$(\psi\circ\phi)(x)=\psi(\phi(x)).$$

证 设 $\phi(\alpha)=(Y,\nu,b)$，$\psi(\phi(\alpha))=(Z,\pi,c)$，则对任一 $z\in Z$，有

$$\begin{aligned}c(Z)&=((\exists y\in Y)\psi(y)=z,\ b(y))\\&=((\exists y\in Y)\psi(y)=z,\ ((\exists x\in X)\phi(x)=y,\ a(x)))\\&=((\exists y\in Y,\ \exists x\in X)\psi(y)=z,\ \phi(x)=y,\ a(x))\\&=((\exists x\in X)\psi(\phi(x))=z,\ a(x)),\end{aligned}$$

这就证明了$\psi(\phi(x))=(Z,\pi,c)$的确是在映射$\psi\circ\phi$之下 $\alpha=(X,\mu,a)$在 Z 中的像。

定义 4.2.3 若 $\alpha=(X,\mu,a)$和 $\beta=(Y,\nu,b)$是两个不分明集，则 α 与 β 的乘积不分明集(Product fuzzy set)$\alpha\times\beta$

为不分明集$(X\times Y,\mu\times\nu,a\wedge b)$，其中$X\times Y$是基本空间$X$和$Y$的笛卡尔直积，$(\mu\times\nu)(x,y)=\min(\mu(x),\nu(y))$，且$(a\wedge b)(x,y)=a(x)\wedge b(y)$。

定义 4.2.4 在实直线R上定义的不分明集$\alpha=(R,\mu,a)$，若μ分段连续，则称α为一个不分明数，简称F数.

定义 4.2.5 使得集

$$\{x\in R\mid\mu(x)\geqslant\lambda\}$$

对每一个$\lambda\in[0,1]$是凸集的不分明数，这种集我们称为单峰的(unimadal)。

一个单峰的不分明数，其隶属函数亦称为单峰的。

若α和β是两个不分明数，则$\alpha+\beta$和$\alpha\cdot\beta$将分别表示在映射$(x,y)\to x+y$和$(x,y)\to xy$之下的直积不分明集$\alpha\times\beta$在R中的像.

定义 4.2.6 一个不分明集(X,μ,a)若存在一个元素$x\in X$，使得$\mu(x)=1$，则称该不分明集为正态的(normal).

4.2.2 不分明变量

在1978年S.Nahmias定义了不分明变量(SF变量)。认为SF变量X为定义在任意的集Γ上的实值函数。从SF变量可以引进F随机变量。

定义 4.2.7 在集Γ上的一切子集的类上的标度σ是满足下面条件的函数：

(i) $\sigma(\phi)=0$，$\sigma(\Gamma)=1$；

(ii) 对于Γ的任何子集系$\{A_\alpha\}$，

$$\sigma\left(\bigcup_\alpha A_\alpha\right)=\sup_\alpha\sigma(A_\alpha).$$

这个标度σ的概念同Zadeh的可能性分布和F测度的概

念是等价的。由 σ 和 F 变量的定义可以确定 F 变量的隶属函数（SF 变量 X 的隶属函数 $\mu_X: R\to[0,1]$ 由

$$\mu_X(x)=\sigma\{r\in\Gamma: X(r)=x\},\quad x\in R$$

定义）。为了对函数 $g: R\to R$ 得到 SF 变量 X 的 $g(X)$ 的隶属函数，Nahmias 证实了 Zadeh 的扩张原理：

$$\mu_{g(X)}(t)=\sup_{u: g(u)=t}\mu_X(u).$$

例（i）　$\mu_{aX}(t)=\mu_X(t/a)$，对 $a\neq 0$，一切 t。

（ii）　$\mu_{X^2}(t)=\mu_X(\sqrt{t})\vee\mu_X(-\sqrt{t})$，$t\geqslant 0$。

两个 SF 变量 X,Y 无关是指对一切 x,y 有

$$\sigma(\{X=x\}\cap\{Y=y\})=\sigma\{X=x\}\wedge\sigma\{Y=y\}.$$

这个概念类似随机变量独立性的概念。一些 SF 变量称为相互无关是指它的任何一部分（有限个）交的标度能通过每项标度的最小标度计算出来（见 Nahmias 1978）。

定理 4.2.2　若 X，Y 是 SF 变量，则

$$\mu_{X+Y}(t)=\sup_x\sigma(\{X=x\}\cap\{Y=t-x\}).$$

又设 X，Y 无关，则

$$\mu_{X+Y}(t)=\sup_x[\mu_X(x)\wedge\mu_Y(t-x)].$$

其中 $X+Y$ 是由

$$(X+Y)(r)=X(r)+Y(r),\quad r\in\Gamma$$

来定义的。

按照 SF 变量的定义，一切 SF 变量之集 R^Γ 是 R 上的向量空间。从而可以定义数量积为通常的数量积。而 F 变量的积与商可通过实数对之间的二元运算 $*$ 来定义：$X*Y$ 由 $(X*Y)(r)=X(r)*Y(r)$ 来定义，其隶属函数在 X，Y 无关时

可证明为

$$\mu_{Z*Y}(t)=\sup_{X*Y=t}[\mu_X(x)\wedge\mu_Y(y)]。$$

SF变量有一个重要的特性是所谓的凸性。一个 SF 变量 X 为凸的是说 X 的隶属函数是半凹(quasi-concave)。即对一切 a，$b\in R$ 和 $0\leqslant\lambda\leqslant1$，有

$$\mu_X(\lambda a+(1-\lambda)b)\geqslant\mu_X(a)\wedge\mu_X(b)。$$

遵从惯例，仍称 μ_X 是凸的。Dubois 和 Prade 称一个凸隶属函数为一个 F 数.

1975 年 Chang 证了如果 f：$R^n\to R$ 是一个连续函数，且 X_1，…，X_n 是无关的 SF变量，则(X_1，…，X_n)仍为凸 SF 变量。

凸隶属函数类包括所有的脆弱集及一切单调函数。也包括一切 $N(a, b)$函数，这里 $\mu(x)=e^{-(x-a)^2/b^2}$，$x\in R$，$a\in R$，$b>0$。

W. E. Stein 和 K. Talti 研究了凸 SF 随机变量.他认为 Zadeh 的 F 集定义有毛病，从而追随 S. Nahimas 的观点研究这一问题。下面我们介绍 F 随机变量。

§3　F 随机变量

不分明随机变量，简称为 F随机变量。有两种不同的定义办法,分别由Kwakernaak 及 Stein 等人所引进。我们在 4.3.1 及 4.3.2 中给以介绍。

4.3.1　F随机变量是一个取值为F数的随机变量

设(Ω，$\mathscr{F}$，P)为经典概率空间，U 为定义于其上的随机变量，w_i($i\in J$，J 为一有限的或可数集)是窗集，即$w_i\subset$

R，$w_i\cap w_j=\varnothing$，$(i\neq j)$，且$\bigcup_{i\in J}w_i=R$.通过窗集可"看出"随机变量ξ的含义是对每个ω，仅能确定对某些$i\in J$，是否$U(\omega)\in w_i$. 又设w_i的特征函数为I_i：$R\to[0,1]$，而用S表示一切分段连续函数映射$R\to[0,1]$的空间. 如前4.2.1中的讨论，随机变量U为由$\omega\overset{X}{\to}X_\omega$给定的映射$X$：$\Omega\to S$所确定。当且仅当$U(\omega)\in w_i$时，有$X_\omega=I_i$。这指明与每个$\omega\in\Omega$相联系的$U(\omega)$不是一个实数，但是$X_\omega\in S$。因此，映射$X$：$\Omega\to S$应当是$F$随机变量的一个特殊类型。我们称$U$为$F$随机变量的原型。记一切$F$数的集为$\boldsymbol{F}$。

定义 4.3.1 称映射ξ：$\Omega\to\boldsymbol{F}$为F随机变量。

在定义中，注意$\xi(\omega)=(R,X_\omega,a_\omega)$，$X_\omega\in S$，$A_\omega$：$R\to M$。而由$\omega\overset{X}{\to}X_\omega$所确定的映射$X$：$\Omega\to S$，要求对每个$\mu\in[0,1]$，$U_\mu^*$和$U_\mu^{**}$是$(\Omega,\mathscr{F},P)$上的有限实随机变量，满足

$$(\forall\omega\in\Omega)\quad X_\omega(U_\mu^*(\omega))\geqslant\mu,$$
$$X_\omega(U_\mu^{**}(\omega))\geqslant\mu.$$

其中U_μ^*，U_μ^{**}分别由下式定义：

$$U_\mu^*(\omega)=\inf\{x\in R\,|\,X_\omega(x)\geqslant\mu\};$$
$$U_\mu^{**}(\omega)=\mathrm{Sup}\{x\in R\,|\,X_\omega(x)\geqslant\mu\}.$$

最后，对每个$\omega\in\Omega$，$x\in R$，$a_\omega(x)$是语句"在点ω最初假定的值x"。

F随机变量ξ可解释得更为清楚些。设概率空间$(\widetilde{\Omega},\widetilde{\mathscr{F}},\widetilde{P})$比$(\Omega,\mathscr{F},P)$更广泛。又设$\tilde{\sigma}$是$\widetilde{\mathscr{F}}$的子$\sigma$代数，在$\widetilde{\Omega}$上定义一个等价关系～为

$$\tilde{\omega}_1\sim\tilde{\omega}_2\Leftrightarrow[\forall\tilde{A}\in\tilde{\sigma}),\ \tilde{\omega}_1\in\tilde{A}\Leftrightarrow\omega_2\in\tilde{A}].$$

由上我们由问题得到一个 F 随机变量的一切可能原型组成的$(\widetilde{\mathscr{X}},\ t(b(\cdot)),b)$。为了避免符号混乱，以后不分明集记为$\boldsymbol{X}=(\widetilde{\mathscr{X}},X)$。我们沿用习惯的叫法，称 $\boldsymbol{X}$ 为一个 F随机变量。换句话说，称映射 ξ 为一个 F随机变量。在以后诸节内称其为 KF 随机变量。

当对每个 $\omega\in\Omega$，存在一个 $x\in R$，使 $X_\omega(x)=1$ 的 F 随机变量称为是正态的。在 Kwakernaak 的研究中仅考虑正态情形。

定义 4.3.2 若在定义 4.3.1 中 Ω 为离散的可数集，则称 $\boldsymbol{X}$ 为离散的 F 随机变量。

当讨论单个 F 随机变量时，可让 $\Omega=N$，N为自然数集，$\mathscr{F}$取为 N 的子集的 σ 代数，并记 $P(\{i\})=p_i$，$i\in N$。且

$$i\overset{X}{\rightarrow}X^i,\ \text{对一切}\ i\in N.$$

可见，离散的单一F随机变量$\boldsymbol{X}$ 实质是由一列数对$(p_i,\ X^i)$ $(i=1,2,\cdots)$来刻划的。其中$\sum\limits_i p_i=1$，X^i 是隶属函数，表示由 F 随机变量决定的 F 值。

在意见询问的例子中（见第二章例 2.2.1），其结果为一离散 F 随机变量。概率 $p_1=0.4$，$p_2=0.5$，$p_3=0.1$，而X^1，X^2，X^3分别是“很暖”、“暖”及“不评价”的隶属函数。

离散 F 随机变量 $\boldsymbol{X}$，其原型关于 $\sigma(\boldsymbol{X})\otimes\mathscr{F}'$ 可测，且是定义在$N\times\Omega'$上的随机变量$\widetilde{U}$。将 $\widetilde{U}$ 在$(i,\omega')\in N\times\Omega'$的值记为$\widetilde{U}_i(\omega')$。$\widetilde{U}$ 在 F 集 $\boldsymbol{X}=(\mathscr{X},x)$上的隶属度如下：

$$\inf_{i\in N}\inf_{\omega\in\Omega}X^i(\widetilde{U}_i(\omega')).$$

对于辅助概率空间$(\Omega',\ \mathscr{F}',\ P')$可如下理解：最初的

代数 $\tilde{\mathfrak{F}}\subset\widetilde{\mathscr{F}}$ 而论，它由一切形如 $\tilde{A}=A\times\Omega'$，$A\in\mathscr{F}$ 的柱集所组成。

设 ξ 是一个 F 随机变量，使得

$$\omega\xrightarrow{\xi}(R,X_\omega,a_\omega).$$

用 $\sigma(X)$ 记由随机变量 U_μ^* 及 U_μ^{**} 生成的 Ω 子集的 σ 代数，其中 $\mu\in[0,1]$。为简单计，把 $\sigma(X)$ 作为由 X 生成的 σ 代数。现在我们来说明一个事实，概率空间 $(\Omega,\ \mathscr{F},\ P)$ 上的 ξ，按它的原型，是 $(\tilde{\Omega},\ \widetilde{\mathscr{F}},\ \tilde{P})$ 上的随机变量。ξ 表现为一个不分明约化。对 $(\tilde{\Omega},\ \widetilde{\mathscr{F}},\ \tilde{P})$ 要求任何一个 ξ 的原型 $\tilde{U}$ 关于 $\sigma(X)\otimes\mathscr{F}'$ 可测，其意思是在 Ω 中 $\tilde{U}$ 与有用的不分明信息相一致。

因此，ξ 的一切可能的原型之集 $\mathscr{X}$ 是定义在 $(\tilde{\Omega},\ \widetilde{\mathscr{F}},\ \tilde{P})$ 上的一切随机变量的集合，它关于 $\sigma(X)\otimes\mathscr{F}'$ 可测.对任何的 $\tilde{U}\in\widetilde{\mathscr{X}}$，可接受的是由语句 $b(\tilde{U})$ 的真值给定的原型，这里

$$b(\tilde{U})=(\tilde{U}\text{是 }\xi\text{ 的原型})$$

$$=\bigwedge_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}}(\text{在点}(\omega,\omega')\text{的原型值}\tilde{U}(\omega,\omega'))$$

$$=\bigwedge_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}}a_\omega(\tilde{U}(\omega,\omega'))$$

所以，可接受的 $\tilde{U}$ 是由下式给定的原型：

$$t(b(\tilde{U}))=\inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}}t(a_\omega(\tilde{U}(\omega,\omega')))$$

$$=\inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}}X_\omega(\tilde{U}(\omega,\omega')).$$

则我们可取 Ω 为在关系～下的商集$\Omega=\tilde{\Omega}/\sim$. 即在～关系下将$\tilde{\Omega}$ 分为相互不相交的等价类。为了定义 $\mathscr{F}$ 和 P，令 c 为正规映射c：$\tilde{\Omega}\to\Omega$，而 $\mathscr{F}$ 为 Ω 的子集类，它在 c 映射下的原像属于$\tilde{\mathscr{F}}$。测度 P 定义为：若 $A\in\mathscr{F}$，则 $P(A)=\tilde{P}[c^{-1}(A)]$. 我们称$(\Omega,\ \mathscr{F},\ P)$为由 $\tilde{\sigma}\subset\tilde{\mathscr{F}}$ 生成的$(\tilde{\Omega},\ \tilde{\mathscr{F}},\ \tilde{P})$的约化。

设 $\tilde{U}$ 是任一个定义在$(\tilde{\Omega},\ \tilde{\mathscr{F}},\ \tilde{P})$上 $\tilde{\sigma}$ 可测的随机变量，则可得对 $\omega\in\Omega$，$\tilde{U}$ 在集 $c^{-1}(\omega)$上为常量。事实上，因为若$\tilde{\omega}_1\in c^{-1}(\omega)$，$\tilde{\omega}_2\in c^{-1}(\omega)$。则 $c(\tilde{\omega}_1)=c(\tilde{\omega}_2)$。因此$\tilde{\omega}_1\sim\tilde{\omega}_2$。若 $\tilde{U}(\omega_1)=x$，那么 $\tilde{\omega}_1\in\{\tilde{\omega}\,|\,\tilde{U}(\tilde{\omega})=x\}$。于是由$\tilde{\omega}_1\sim\tilde{\omega}_2$有$\tilde{\omega}_2\in\{\tilde{\omega}\,|\,\tilde{U}(\omega)=x\}$.所以$\tilde{U}(\omega_2)=x=\tilde{U}(\tilde{\omega}_1)$。故证得 $\tilde{U}$ 在$c^{-1}(\omega)$上为常量。

现在，在$(\Omega,\ \mathscr{F},\ P)$上用$U(\omega)=\tilde{U}(c^{-1}(\omega))$定义随机变量 U，显然 $\tilde{U}$ 和 U 有相同的概率分布。给定关于 $\tilde{\sigma}$ 可测的任一个随机变量集 $\tilde{U}_1$，…，$\tilde{U}_n$，定义在$(\tilde{\Omega},\ \tilde{\mathscr{F}},\ \tilde{P})$上，则可得一个定义在$(\Omega,\ \mathscr{F},\ P)$上的随机变量集$U_1, U_2,\cdots, U_n$，并且 U_1，…，U_n 与 $\tilde{U}_1$，…，$\tilde{U}_n$ 有同样的概率分布。

可见，给定概率空间$(\Omega,\ \mathscr{F},\ P)$，就存在一个概率空间$(\tilde{\Omega},\tilde{\mathscr{F}},\ \tilde{P})$，使$(\Omega,\ \mathscr{F},\ P)$为其约化。设$(\Omega,\ \mathscr{F},P)$为最小的。从而它是它自己的约化(关于$\mathscr{F}$)。因此，我们引进一个辅助的概率空间$(\Omega',\ \mathscr{F}',\ P')$，假定任何有限个随机变量集恰有给定的分布。目前设$(\tilde{\Omega},\ \tilde{\mathscr{F}},\ \tilde{P})=(\Omega\times\Omega',\mathscr{F}\otimes\mathscr{F}',\ P\otimes P')$，这里$\mathscr{F}\otimes\mathscr{F}'$是包含一切形如 $A\times A'$，$A\in\mathscr{F}$，$A'\in\mathscr{F}'$的集的最小 σ 代数，并且 $P\otimes P'$ 是$\mathscr{F}\otimes\mathscr{F}'$ 上的直积测度。显然$(\Omega,\ \mathscr{F},\ P)$是$(\tilde{\Omega},\ \tilde{\mathscr{F}},\ \tilde{P})$的约化，就 σ

样本空间 N 作为该地被分成不相交的团体数，相应于第 i 个团体 F 随机变量的值 X^i 认为是该团体交出的意见。因此，第 i 个团体要求一个实数 u，作为该团体(在我们例子中涉及到温度)可接受的意见是 $X^i(u)$。引入辅助概率空间的目的，在于将每个团体的意见区分开来。对于固定的 i，随机变量 $\widetilde{U}_i(\omega')$，$\omega'\in\Omega'$，表示第 i 团体范围内关于意见的划分。采纳 $\widetilde{U}_i(\omega')$ 实际相当于第 i 个团体的意见是 $\inf\limits_{\omega'\in\Omega'} X^i(\widetilde{U}_i(\omega'))$. 采纳 $\widetilde{U}$ 相当于对一切团体的意见由上面的式子给出.

4.3.2 SF 变量的概念可以扩展到 F 随机变量

设 Ω 为样本空间，$\Omega=(\omega)$。概率测度定义在 Ω 子集的一个 σ 代数上，R^Γ 表示定义在任意集 Γ 上的一切实值函数。称 Z：$\Omega\to R^\Gamma$ 为 F 随机变量。按照 F 变量定义，$Z(\omega),\omega\in\Omega$，是一个 F 变量。

现在我们仅考虑取有限值的 F 随机变量。设 X_1，$\cdots,X_n$ 是 F 变量，E_1，$\cdots$，E_n 为 Ω 的一个划分．称 $Z(\omega)=\sum\limits_{i=1}^{n} I_{E_i}(\omega)X_i$ 为一取有限个“值”(每个是 F 变量)的 F 随机变量．因此，Z 依概率 $P(E_i)=P_i$ 取值 $Z(\omega)$。

这种 F 随机变量今后称为 SF 随机变量．

§4　F 随机变量的数学期望与方差

在概率论中经常应用随机变量的数字特征，它作为研究分布函数特征的工具而在初等概率论中占有重要的地位。早在古典概率的崛起时期，数学期望的概念已隐含于文章之中，经过一段相当长的时间之后，方才有随机变量的数学期

望这一名称。其本身的实际意义乃是随机变量的平均值，表明了随机变量取值的位置特征，这些都是众所周知的。

4.4.1 经典随机变量的数学期望

当我们知道随机变量 ξ 的分布函数 $F_\xi(x)$ 以后，借助于勒贝格积分，可以直接给出数学期望 $E\xi$ 的定义。当然，我们亦可先定义随机变量矩的概念，而后以数学期望为其特例而得出，这里按前一种办法给出定义。

定义 4.4.1 设随机变量 ξ 的分布函数为 $F_\xi(x)$，若积分 $\int_{-\infty}^{\infty}|x|dF_\xi(x)$ 存在，则称积分

$$\int_{-\infty}^{\infty}xdF(x) \tag{4.4.1}$$

为 ξ 的数学期望(均值或期望)，记作 $E\xi$。

如果我们从概率公理体系出发，我们亦可定义期望。

定义 4.4.2 设 $\xi(\omega)$ 是定义概率空间$(\Omega, \mathscr{F}, P)$上的随机变量，$\omega\in\Omega$。若积分 $\int_\Omega|\xi(\omega)|dP(\omega)$ 存在，则称

$$\int_\Omega\xi(\omega)dP(\omega) \tag{4.4.2}$$

为随机变量 ξ 的数学期望(或均值或期望)，记作 $E\xi$。

这两个定义是等价的. 即

$$\text{定义 }1\Leftrightarrow\text{定义 }2,$$

且

$$\int_\Omega\xi(\omega)dP(\omega)=\int_{-\infty}^{\infty}xdF_\xi(x),$$

等式一端存在，另一端必存在。

这个等式给出了由定义 2 定义的期望 $E\xi$ 的便于计算的公式。而其中积分理解为分布函数 $F_\xi(x)$产生的 Lebesque-

Stieltjes 测度的 Lebesque-Stieltjes 积分。如此，$E\xi$表为数轴上的积分。

4.4.2 F 事件的均值与方差

(1) 1968年 Zadeh 给出如下定义

定义4.4.3 F 事件 $\underset{\sim}{A}$的均值和方差分别为

$$m(\underset{\sim}{A})=\frac{1}{P(\underset{\sim}{A})}\int_{R^n}x\underset{\sim}{A}(x)dP, \tag{4.4.3}$$

$$D^2(\underset{\sim}{A})=\frac{1}{P(\underset{\sim}{A})}\int_{R^n}(x-m(\underset{\sim}{A}))^2\underset{\sim}{A}(x)dP. \tag{4.4.4}$$

(2) 1974 年 Nanmias 给出另一种形式。设$(X, G(X), P)$为一概率空间，

定义 4.4.4 设 $h(x)\in G(x)$，称

$$m(h(x))=\frac{1}{P(h(x))}\int_X xh(x)dP \tag{4.4.5}$$

$$D^2(h(x))=\frac{1}{P(h(x))}\int_X[x-m(h(x))]^2h(x)dP \tag{4.4.6}$$

分别为 $h(x)$的数学期望及方差。

定义 4.4.5 设不分明事件 $\underset{\sim}{A}$ 的隶属函数为 $h(x)$，则称 $m(h(x))$ 及 $D^2(h(x))$ 为 $\underset{\sim}{A}$ 的数学期望及方差，分别记作 $m(\underset{\sim}{A})$及 $D^2(\underset{\sim}{A})$。

这种定义，实际上是把期望与方差作为概率分布的泛函来加以定义的。

(3) 设$(\Omega, \mathscr{B}, \mu)$为 F 测度空间，其中 Ω 是样本空间，$\mathscr{B}$是一个σ代数，$w(\cdot)$为 F 测度。又设 χ 为 F 集合 $\underset{\sim}{A}\subset\Omega$ 的隶属函数，关于$\mathscr{B}$可测。

定义 4.4.6(A. Kandel) 在 F 集合 $\underset{\sim}{A}$ 上 χ关于测度w

的 F 期望值 FEV 由下式确定

$$FEV=\sup_{\alpha\in[0,1]}[\min(\alpha,w(\underset{\sim}{A}_\alpha))] \quad (4.4.7)$$

其中 $\underset{\sim}{A}_\alpha=\{x|\chi(x)\geqslant\alpha\}\subseteq\underset{\sim}{A}$。

引理 4.4.1 对任意的 α_1，$\alpha_2\in[0,1]$，有

$$\min[\alpha_1,w(\underset{\sim}{A}_1)]\leqslant\max[\alpha_2,w(\underset{\sim}{A}_{\alpha_2})] \quad (4.4.8)$$

证 当 $\alpha_1\leqslant\alpha_2$ 时，有 $\min[\alpha_1, w(\underset{\sim}{A}_{\alpha_1})]\leqslant\alpha_1\leqslant\alpha_2\leqslant\max[\alpha_2, w(\underset{\sim}{A}_{\alpha_2})]$。而当 $\alpha_2<\alpha_1$ 时有 $w(\underset{\sim}{A}_{\alpha_1})\leqslant w(\underset{\sim}{A}_{\alpha_2})$，从而 $\min[\alpha_1,w(\underset{\sim}{A}_{\alpha_1})]\leqslant w(\underset{\sim}{A}_{\alpha_1})$

引理 4.4.2 $\sup\limits_{\alpha\in[0,1]}\{\min[\alpha,w(\underset{\sim}{A}_\alpha)]\}\leqslant$

$$\operatorname{Inf}_{\alpha\in[0,1]}\{\max[\alpha,w(\underset{\sim}{A}_\alpha)]\}.$$

证 由引理 4.4.1 $\{\min[\alpha, w(\underset{\sim}{A}_\alpha)]\}$ 的每个数都比 $\{\max[\alpha, w(\underset{\sim}{A}_\alpha)]\}$ 的每个数小，得证。

引理 4.4.3 不存在实数 S，使

$$\sup_{\alpha\in[0,1]}\{\min[\alpha, w(\underset{\sim}{A}_\alpha)]\}<S$$

$$<\operatorname{Inf}_{\alpha\in[0,1]}\{\max[\alpha, w(\underset{\sim}{A}_\alpha)]\}. \quad (4.4.9)$$

证 假如这样的 S 存在。当 $w(\underset{\sim}{A}_S)\leqslant S$ 时，

$$S<\operatorname{Inf}_{\alpha\in[0,1]}\{\max[\alpha, w(\underset{\sim}{A}_\alpha)]\}\leqslant\max[S,w(\underset{\sim}{A}_S)]=S,$$

而当 $w(\underset{\sim}{A}_S)>S$ 时，有

$$S=\min[S, w(\underset{\sim}{A}_S)]\leqslant\operatorname{Sup}_{\alpha\in[0,1]}\{\min[\alpha,$$

$$w(\underset{\sim}{A}_\alpha)]\}<S$$

这两种情形均导致矛盾。从而证得引理。

引理 4.4.4 $\sup\limits_{\alpha\in[0,1]}\{\min[\alpha, w(\underset{\sim}{A}_\alpha)]\}$

$< \operatorname*{Inf}_{\alpha\in[0,1]} \{\max[\alpha, \omega(\underset{\sim}{A}_\alpha)]\}$不真。

证 若它成立，则存在一个实数介于其间，这与引理4.4.3相矛盾。

由引理 4.4.2 至 4.4.4 得如下定理：

定理 4.4.1 $\sup_{\alpha\in[0,1]} \{\min[\alpha, w(\underset{\sim}{A}_\alpha)]\}$

$$= \operatorname*{Inf}_{\alpha\in[0,1]} \{\max[\alpha, w(\underset{\sim}{A}_\alpha)]\}. \qquad (4.4.10)$$

下面我们来讨论集中趋势的度量与 F 期望值之间的关系。

平均数是一组数据的代表。它处于按大小排列的一组数据的中心位置，被认为是集中趋势的测度，它有多种类型。公认的有算术平均数、中位数、众数、几何平均数和调和平均数。

设一组有限数据，有 $n+1$ 个不同的隶属等级$0\leqslant\alpha_1\leqslant\alpha_2\leqslant\cdots\leqslant\alpha_{n+1}\leqslant1$。它蕴含 F 测度 $w(\underset{\sim}{A}_\alpha)$的 n 个不同的水平，其中不考虑 0 与 1。显然集合

$$\{\alpha_i\}_{i=1}^{n+1} \quad 和 \quad \{\omega_j(A_\alpha)\}_{j=1}^{n} \qquad (4.4.11)$$

递增而且在每个集中不能任意交换。因此这两个集合的并有$2n+1$个元素。

定理 4.4.2 如上所述的 $2n+1$ 个数的中位数，对有限的 n 由 FEV 给定。

4.4.3 F 随机变量的数学期望

(1) 设 $\boldsymbol{X}$ 是一个 KF 随机变量

定义4.4.7 $\boldsymbol{X}$ 的期望是一个 F 数，它是 F 集$\boldsymbol{X}=(\widetilde{\mathscr{X}}, X)$在映射 E：$\widetilde{\mathscr{X}}\rightarrow R$ 下的像，使得

$$\widetilde{U} \xrightarrow{E} E\widetilde{U}. \tag{4.4.12}$$

注意，E表示通常的数学期望。F 数 $E\boldsymbol{X}$ 的隶属函数记为$(E\boldsymbol{X})$，因而 $E\boldsymbol{X}=(R,(E\boldsymbol{X}))$。其中

$$(E\boldsymbol{X})(x)=\sup_{\widetilde{U}\in\widetilde{\mathscr{X}}:E\widetilde{U}=x}\ \inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}} X_\omega(\widetilde{U}(\omega,\omega')),$$

$$x\in R. \tag{4.4.13}$$

当 $\boldsymbol{X}$ 为离散 KF 随机变量时，$E\boldsymbol{X}$ 的隶属函数$(E\boldsymbol{X})$ 为

$$\sup_{\widetilde{U}\in\widetilde{\mathscr{X}}:E\widetilde{U}=x}\ \inf_{i\in N}\ \inf_{\omega'\in\Omega'} X^i(\widetilde{U}_i(\omega')). \tag{4.4.14}$$

为了计算 KF 随机变量的(不分明) 期望，就应寻找相应于一给定 F 集的水平集族。如果这些水平集是已知的，该F集就完全被刻划出来。假如(M,N)是一个 F 集，定义在基本空间 M 上，其中 m：$M\to[0,1]$ 为隶属函数。那么对每个$\mu\in[0,1]$，集$C_\mu=\{\alpha\in M|m(\alpha)\geqslant\mu\}$为 μ 水平集。给定水平集族 $C=(C_\mu)$。隶属函数 m 可从下面公式得到：

$$m(\alpha)=\sup\{\mu\in[0,1]|\alpha\in C_\mu\},\ \alpha\in M \tag{4.4.15}$$

设离散 KF 随机变量$\boldsymbol{X}$用 $(p_i,\ X^i)$，$i=1,2\cdots$来描述。现介绍确定 $E\boldsymbol{X}$ 的水平集的计算规则如下：

步骤 1　选择水平 $\mu\in[0,1]$，

步骤 2　对每个 $i=1,2\cdots$确定数

$$u_i^*(\mu)=\inf\{x\in R|X^i(x)\geqslant\mu\},$$

$$u_i^{**}(\mu)=\sup\{x\in R|X_i(x)\geqslant\mu\}$$

步骤 3　确定水平集 $C_\mu=\{x\in R|(E\boldsymbol{X})(x)\geqslant\mu\}$

为

$$C_\mu=\left[\sum_i p_i u_i^*(\mu),\sum_i p_i u_i^{**}(\mu)\right].$$

步骤 4 对足够多的 μ 值重复步骤 1 至 3。

步骤 5 确定水平集 $D_0=\{x\in R|(E\boldsymbol{X})(x)>0\}$为

$$D_0=\left[\sum_i p_i u_i^0,\ \sum_i p_i u_i^{00}\right]$$

其中

$$u_i^0=\inf\{x\in R|X^i(x)>0\},$$
$$u_i^{00}=\sup\{x\in R|X^i(x)>0\}.$$

证 步骤 3 与 5 可由下面定理 4.4.4 得知，步骤 5 是为了限定隶属函数$(E\boldsymbol{X})$的柱集D_0，而$(E\boldsymbol{X})$的形状可由步骤3所得水平集借助(4.4.15)公式推出。

在列出该规则的步骤时，我们总认为 $\boldsymbol{X}$ 满足条件：

$$u_i^*(\mu)\triangleq\inf\{x\in R|X^i(x)\geqslant\mu\}>-\infty,$$
$$u_i^{**}(\mu)\triangleq\sup\{x\in R|X^i(x)\geqslant\mu\}<+\infty,$$

其中

$$X^i(u_i^*(\mu))\geqslant\mu,\qquad X^i(u_i^{**}(\mu))\geqslant\mu.$$

前两个条件说明对每个 $i\in N$，函数X^i 有有限的柱集。如果每个X^i 为上半连续的，即对每个$\mu\in[0,1]$,集$\{x\in R|X^i(x)\geqslant\mu\}$为闭集,后两个条件成立,但实际情形是半连续条件常常不出现。

例 4.4.1 在我们谈及的意见询问问题中，$p_1=0.4$，$p_2=0.5$，$p_3=0.1$，$p_k=0$，$(K\geqslant4)$。而 X^1、X^2 和 X^3 是“很暖”、“暖”和“不评价”的隶属函数(见图 4.4.1)。由规则容易算出$E\boldsymbol{X}$的水平集，再由(4.4.15)公式，算出$(E\boldsymbol{X})$，如图 4.4.2 所示。注意平均温度在 27°C 和 27.5°C 之间是完全合理的，如果温度减小至 22°C或增加到 32.5°C 它减小到 0。

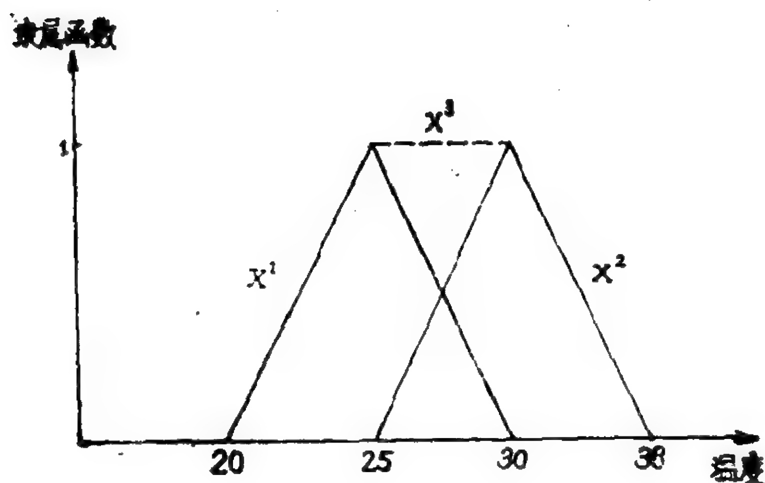


图 4.4.1 X^1, X^2, X^3 之图形

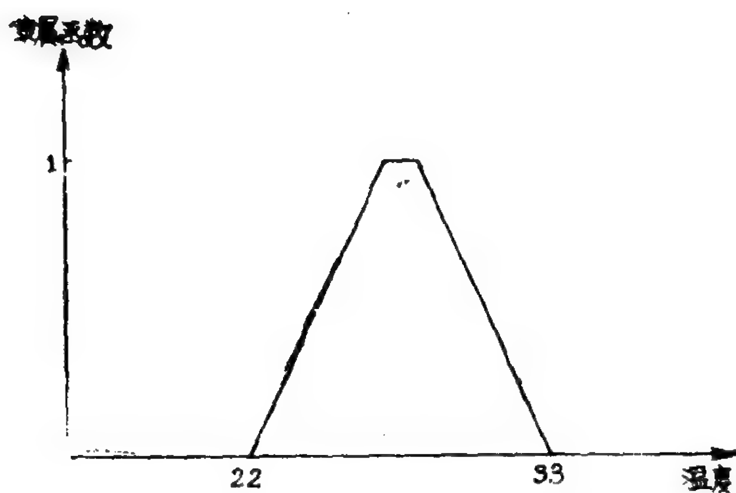


图 4.4.2 $E\boldsymbol{X}$ 的隶属函数($E\boldsymbol{X}$)

一般的说，关于 $E\boldsymbol{X}$ 有如下定理. 设 $\mathscr{X}$ 是定义在(Ω, $\mathscr{F}$, P)上的 $\sigma(X)$可测随机变量的集合.

定理 4.4.3 若 $\boldsymbol{X}$ 是单峰的，即 X_ω 为单峰的，则

$$(E\boldsymbol{X})(x)=\sup_{U\in\mathscr{X}:EU=x}\ \inf_{\omega\in\Omega} X_\omega(U(\omega)),\quad x\in R \tag{4.4.16}$$

证 设 $\varepsilon>0$ 是任意一个正实数。则总存在一个随机变量$\widetilde{U}^*\in\widetilde{\mathscr{X}}$，且 $E\widetilde{U}^*=x$，它在上式中按精度 ε 达到上确界，即

$$(E\boldsymbol{X})(x)=\inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}} X_\omega(\widetilde{U}^*(\omega,\omega'))+\varepsilon.$$

定义随机变量

$$V^*(\omega)=\int\widetilde{U}^*(\omega,\omega')dP'(\omega'),\quad \omega\in\Omega.$$

显然 $V^*\in\mathscr{X}$，且 $EV^*=x$。由 X_ω 的单峰性，故对一切 $\omega\in\Omega$，有

$$\inf_{\omega'\in\Omega'} X_\omega(\widetilde{U}^*(\omega,\ \omega'))\leqslant X_\omega(V^*(\omega)),$$

因此

$$(E\boldsymbol{X})(x)\leqslant\inf_{\omega\in\Omega}X_\omega(V^*(\omega))+\varepsilon$$

$$\leqslant\sup_{V\in\mathscr{X}:EV=x}\ \inf_{\omega\in\Omega}X_\omega(V(\omega))+\varepsilon$$

另一方面，由(4.4.13)式限制 $\widetilde{U}$，使 $\widetilde{U}(\omega,\omega')=V(\omega)$，$V\in\mathscr{X}$，显然

$$(E\boldsymbol{X})(x)\geqslant\sup_{V\in\mathscr{X}:EV=x}\ \inf_{\omega\in\Omega}X_\omega(V(\omega)),$$

将它们结合起来，有

$$\sup_{V\in\mathscr{X}:EV=x}\ \inf_{\omega\in\Omega}X_\omega(V(\omega))\leqslant(E\boldsymbol{X})(x)$$

$$\leqslant\sup_{V\in\mathscr{X}:EV=x}\ \inf_{\omega\in\Omega}X_\omega(V(\omega))+\varepsilon$$

由于 ε 的任意性，证得定理。

这个定理说明，若 $\boldsymbol{X}$ 是单峰的，可直接由$(\Omega,\mathscr{F},P)$确定 $E\boldsymbol{X}$。

对于一个给定的KF 随机变量 $\boldsymbol{X}$，定义：$\forall\omega\in\Omega$，

$$\overline{X}_\omega(x) = \sup_{U, \cdot \in R; u \leqslant x \leqslant v} \min[X_\omega(u), X_\omega(v)], \quad x \in R,$$

容易看出 $\overline{X}_\omega$ 是单峰的。假如 X_ω 是单峰的，则 $\overline{X}_\omega = X_\omega$。我们定义可约的 KF 随机变量 $\overline{\boldsymbol{X}} = (\mathscr{H}, \overline{X})$，即$\mathscr{H}$为定义在$(\Omega, \mathscr{F}, P)$上一切 $\sigma(X)$可测的随机变量的集合。则我们可以证明

定理 4.4.4 $E\boldsymbol{X} = E\overline{\boldsymbol{X}}$。

它说明对于单峰情形求 $E\boldsymbol{X}$ 没有丢掉精度。而且定理结论可以如下看待：映射 E：$\widetilde{\mathscr{H}} \to R$ 可由下面两个映射合成。令 $\widetilde{U} \in \widetilde{\mathscr{H}}$，则由

$$E\widetilde{U} = \int dP(\omega) \int \widetilde{U}(\omega, \omega') dP'(\omega') \tag{4.4.17}$$

可知，$E\widetilde{U}$ 为映射(i)和(ii)的合成：

(i) $\widetilde{\mathscr{X}} \to \mathscr{X}$。由 $\widetilde{U} \to \overline{U}$，$\overline{U}(\omega) = \int \widetilde{U}(\omega, \omega') dP'(\omega')$

确定；

(ii) $\mathscr{X} \to R$。由 $\widetilde{U} \to E\overline{U}$ 指定。

在(i)之下 $\boldsymbol{X}$ 的像为$\overline{\boldsymbol{X}}$；在(ii)下$\overline{\boldsymbol{X}}$ 的像为 $E\overline{\boldsymbol{X}}$。由定理 4.2.1，$\boldsymbol{X}$ 在复合映射之下的像恰好是 $E\overline{\boldsymbol{X}}$。

定义 4.4.8 设 $\boldsymbol{X} = (\widetilde{\mathscr{X}}, X)$和 $\boldsymbol{Y} = (\widetilde{\mathscr{Y}}, Y)$是两个 KF 随机变量。称 R 中直积 F 集 $\boldsymbol{X} \times \boldsymbol{Y} = (\widetilde{\mathscr{X}} \times \widetilde{\mathscr{Y}}, X \times Y)$在映射$(\widetilde{U}, \widetilde{V}) \to E\widetilde{U}\widetilde{V}$ 下的像 $E\boldsymbol{XY}$ 为 $\boldsymbol{XY}$ 的 F 期望。

定义的含意是指 $E\boldsymbol{XY}$ 为一个 F 数，$E\boldsymbol{XY} = (R, (E\boldsymbol{XY}))$。其中对一切 $z \in R$，

$$(E\boldsymbol{XY})(z) = \sup_{\substack{\widetilde{U} \in \widetilde{\mathscr{X}}, \widetilde{V} \in \widetilde{\mathscr{Y}}: \\ E\widetilde{U}\widetilde{V} = z}} \inf_{\substack{\omega \in \Omega \\ \omega' \in \Omega'}} [X_\omega(\widetilde{U}(\omega, \omega')),$$

$$Y_\omega(\widetilde{V}(\omega,\omega'))]$$

因为$(\widetilde{U},\widetilde{V})\in\widetilde{\mathscr{X}}\times\widetilde{\mathscr{Y}}$有隶属度(在F集$\boldsymbol{X}\times\boldsymbol{Y}$中)

$$\min[\inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}} X_\omega(\widetilde{U}(\omega,\omega')),\quad \inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}} Y_\omega(\widetilde{V}(\omega,\omega'))]$$

$$=\inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}}\min[X_\omega(\widetilde{U}(\omega,\omega')),Y_\omega(\widetilde{V}(\omega,\omega'))]$$

一个KF随机变量$\boldsymbol{X}$称为是非负的，若对于$x<0$及一切$\omega\in\Omega$，有$X_\omega(x)=0$。由单峰变量$\overline{\boldsymbol{X}}$和$\overline{\boldsymbol{Y}}$分别给定的$KF$随机变量$\boldsymbol{X}=(\widetilde{\mathscr{X}},X)$和$\boldsymbol{Y}=(\widetilde{\mathscr{Y}},Y)$，有如下结论。

定理 4.4.5 若$\boldsymbol{X}$和$\boldsymbol{Y}$是单峰的，则

$$(E\boldsymbol{XY})(z)=\sup_{\substack{U\in\mathscr{X}\\ V\in\mathscr{Y}}:\ EUV=z}\inf_{\omega\in\Omega}\min[\overline{X}_\omega(U(\omega)),\overline{Y}_\omega(\widetilde{V}(\omega))]\qquad(4.4.18)$$

其中$z\in R$。

定理证明较长，故省略。

定义 4.4.9 若由KF随机变量$\boldsymbol{X}_i=(R,X^i)$ $(i=1,\cdots,n)$的X^i所生成的Ω子集的σ代数$\sigma(X^i)$ $(i=1,\cdots,n)$是独立的，则称这些随机变量$\boldsymbol{X}_i$是独立的。

定理 4.4.6 若$\boldsymbol{X}$和$\boldsymbol{Y}$是两个非负的KF随机变量，则

$$E\boldsymbol{XY}=E\boldsymbol{X}\cdot E\boldsymbol{Y}.\qquad(4.4.19)$$

证 不失普遍性，设$\boldsymbol{X}$和$\boldsymbol{Y}$是单峰的。由于$\boldsymbol{X}$与$\boldsymbol{Y}$独立，故任意的随机变量$U\in\mathscr{X}$和$V\in\mathscr{Y}$也是独立的。从而有

$$(E\boldsymbol{XY})(z)=\sup_{\substack{U\in\mathscr{X}\\ V\in\mathscr{Y}}:\ EUV=z}\inf_{\omega\in\Omega}\min[X_\omega(U(\omega)),$$

$$Y_\omega(V(\omega))]$$

$$= \sup_{\substack{U\in\mathscr{X}, V\in\mathscr{Y}, u, v\in R: \\ EU=u, EV=v, \\ uv=z}} \inf_{\omega\in\Omega} \min[X_\omega(U(\omega)), Y(V(\omega))]$$

$$= \sup_{u,v\in R: uv=Z} \sup_{U\in\mathscr{X}, U\in\mathscr{Y}: EU=u, EV=v} \inf_{\omega\in\Omega}$$

$$\min[X_\omega(U(\omega)), Y_\omega(V(\omega))]$$

$$= \sup_{u,v\in R: uv=Z} \min\Big[\sup_{U\in\mathscr{X}: EU=u} \inf_{\omega\in\Omega} X_\omega(U(\omega)),$$

$$\sup_{V\in\mathscr{Y}: EV=v} \inf_{\omega\in\Omega} Y_\omega(V(\omega))\Big]$$

$$= \sup_{u,v\in R, uv=Z} \min[(E\boldsymbol{X})(u), (E\boldsymbol{Y})(v)]$$

$$= (E\boldsymbol{X}E\boldsymbol{Y})(z).$$

(2) 关于 Stein 的 SF 随机变量．数学期望值仅 考虑形如下面的 SF 随机变量：

$$Z(\omega) = \sum_{i=1}^{n} I_{E_i}(\omega) X_i \qquad (4.4.20)$$

其中 X_1，…，X_n 是 F 变量，E_1，…，E_n 是样本空间 Ω 的一个划分，$\omega\in\Omega$，且 Z 依概率 $P(E_i)=p_i$ 取值 X_i。

定义 4.4.10 Z 的数学期望 EZ 为

$$E(Z) = \sum_{i=1}^{n} p_i X_i. \qquad (4.4.21)$$

定理 4.4.7 对一切 α，β，有

$$E(\alpha Z+\beta) = \alpha E(Z)+\beta. \qquad \alpha, \beta\in R.$$

证 $\alpha Z+\beta = \alpha\sum I_{E_i} X_i + \beta = \sum I_{E_i}(\alpha X+\beta)$，所以

$$E(\alpha X+\beta)=\sum p_i(\alpha X_i+\beta)=\alpha\sum p_i X_i+\beta=\alpha E(Z)+\beta.$$

定理 4.4.8 设$\{X_i\}$是无关的 F 变量，则

$$\mu_{E(Z)}(t)=\sup[\mu_{X_1}(x_1)\wedge\cdots\wedge\mu_{X_n}(x_n)] \quad (4.4.22)$$

其中 sup 遍及$(x_1,\cdots,x_n)$且受 $\sum p_i x_i=t$ 约束。

证 ∵ $\mu_{E(Z)}(t)=\sigma(\sum p_i x_i=t)$. 类似于 § 3.1 的证明过程可得。

注 这一定理结论与 KF 随机变量期望中所举的例子结论相一致。

定理 4.4.9 设 SF 随机变量

$$Z=\begin{cases}\boldsymbol{X}, & \text{依概率 } p,\\ \boldsymbol{Y}, & \text{依概率 } q.\end{cases}$$

其中 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 是无关的 F 变量，且 $p+q=1$，则

$$\mu_{E(Z)}(t)=\sup_x[\mu_{\boldsymbol{X}}(x/p)\wedge\mu_{\boldsymbol{Y}}((t-x)/q)] \quad (4.4.23)$$

证 直接由定理 4.4.4 得证。

定理 4.4.10 设 Z_1，Z_2，$\cdots$，Z_n 是 SF 随机变量，α_1，$\cdots$，α_n 是任意实数。则

$$E(\sum\alpha_i Z_i)=\sum\alpha_i E(Z_i). \quad (4.4.24)$$

这个定理显然是定理 4.4.7 的推广。

定理 4.4.11 设 Z，W 是独立 SF 随机变量，则

$$E(ZW)=E(Z)E(W). \quad (4.4.25)$$

证 设 $Z=\sum I_{E_i}X_i$，$W=\sum I_{F_j}Y_j$，其中$\{E_i\}$及$\{F_i\}$是 Ω 的两个划分。且 Z，W 独立的意思是 $P(E_i\cap F_j)=P(E_i)P(F_j)$。故

$$E(ZW)=\sum\sum P(E_i)P(F_j)X_iY_j$$

$$= \sum P(E_i)X_i \cdot \sum P(F_j)Y_j = E(Z)E(W).$$

从以上关于数学期望的定义来看，*KF* 随机变量的期望是一个 *F* 数，而 *SF* 随机变量的期望是一个 *F* 变量，亦是一个 *SF* 随机变量。但都可看成是 *F* 集合。他们出发点(处理办法)不同，但基本结构相似，虽然 Kwakernaak 的理论中期望不保持线性性质。结合 Zadeh，Nguyen，Kandel 和 Byan 等人所给的 *F* 随机变量来看，更好的结论还待进一步发展。

§5 收敛性

设 W，$\{Z_n\}$是 *SF* 随机变量。

定义 4.5.1 设 X_n，X 是 *SF* 随机变量，若对一切 $r\in\Gamma$，有 $X_n(r)\to X(r)$，则称 X_n 收敛于 X，记作 $X_n\to X$。

这个定义与点态收敛是不相同的。例如，考虑 $\Gamma=\{\alpha,\beta,\delta\}$，每个元素有相同的等级 σ。定义 *F* 变量：

$$X_n(\alpha)=1, \qquad X(\alpha)=1.$$

$$X_n(\beta)=2, \qquad X(\beta)=2.$$

$$X_n(\delta)=2+1/n, \qquad X(\delta)=2.$$

因此 $X_n\to X$。现在

$$\lim \mu_{X_n}(t)=\begin{cases}\sigma(\alpha), & \text{若 } t=1\\ \sigma(\beta), & \text{若 } t=2\\ 0, & \text{其它}\end{cases}$$

而

$$\mu_{\lim X_n}(t)=\begin{cases}\sigma(\alpha) & \text{若 } t=1,\\ \sigma(\beta\cup\delta) & \text{若 } t=2,\\ 0 & \text{其它}\end{cases}$$

其中 μ_X 为 X 的隶属函数，定义为

$$\mu_X(x)=\sigma\{r\in\Gamma:\ X(r)=x\},\ x\in R.$$

若 $\sigma(\delta)>\sigma(\beta)$，则 $\lim\mu_{X_n}(t)$ 与 $\mu_{\lim X_n}(t)$不同。

定理 4.5.1 令$\{Z_i\}$是独立的 SF 随机变量。设

$$Z_i=\begin{cases}\boldsymbol{X}, & \text{按概率 } p,\\ \boldsymbol{Y}, & \text{按概率 } q,\end{cases}$$

且 $\bar{Z}_n=(Z_1+\cdots+Z_n)/n$，$E(Z_i)=E(Z)$，$i\in N$。则

$$\bar{Z}_n\to E(Z),\ n\to\infty.$$

证 $\bar{Z}_n=(k/n)\boldsymbol{X}+((n-k)/n)\boldsymbol{Y}$，这里 k 是一个随机变量，表示在开头的 n 次试验中"发生"的数目。由非不分明强大数定律，k/n 对每个 ω 收敛于 p。因此，$\bar{Z}_n$ 对每个 ω收敛于 $p\boldsymbol{X}+q\boldsymbol{Y}$，证完。

这个定律可以看作是 F 大数定律的特殊情形。

引理 4.5.1(Nahmias) 若 $X_1,\cdots,X_n$ 是无关 F 变量，X_i 的隶属函数为 $N(a_i,b_i)$，$b_i>0$，即 $\mu_{X_i}(x)=\exp\{-(x-a_i)^2/b_i^2\}$，则对任何正实数 c_1，$\cdots$，c_n，有 $c_1X_1+\cdots+c_nX_n$ 是一个 F 变量，其隶属函数 $N(\sum c_ia_i,\sum c_ib_i)$。

定理 4.5.2 令$\{Z_i\}$是独立的 SF 随机变量。

$$Z_i=\begin{cases}\boldsymbol{X}_i, & \text{依概率 } p,\\ \boldsymbol{Y}_i, & \text{依概率 } q.\end{cases}$$

其中$\{\boldsymbol{X}_i\}$和$\{\boldsymbol{Y}_i\}$每一个是无关 F 变量序列(并且一切 $\boldsymbol{X}_i$和 $\boldsymbol{Y}_i$ 是无关的)。设 $\boldsymbol{X}_i$ 的隶属函数是$N(1,1)$,而 $\boldsymbol{Y}_i$ 是$N(0,1)$的.则对每个 $\omega\in\Omega$,序列 $\bar{Z}_n(\omega)$的隶属函数收敛于 $E(Z_1)=p\boldsymbol{X}_1+q\boldsymbol{Y}_1$ 的隶属函数。

证 $\bar{Z}_n(\omega)$的隶属函数和 $(k/n)\boldsymbol{X}_1+((n-k)/n)\boldsymbol{Y}_1$ 的相同，这里 k 是一个二项随机变量的值。由引理知 $\boldsymbol{X}_1+\boldsymbol{X}_2$

与 $2\boldsymbol{X}_1$ 有相同的隶属函数，且 $\bar{Z}_n(\omega)$ 有隶属函数 $N(k/n,1)$，当 $n\to\infty$ 时，由于 $N(a,b)$ 的连续性，它收敛于 $N(p,1)$ 的 F 变量，容易验证 $E(Z_1)$ 的隶属函数是 $N(p,1)$。

§6　条件期望与条件概率

条件期望可以从很一般的情形来研究。例如 Kwakernaak 就一子 σ 代数条件下 F 随机变量条件期望的定义及性质进行了探讨。但我们这里仅就其特殊情形来考虑。

设 $\boldsymbol{X}=(\widetilde{\mathscr{X}},X)$ 和 $\boldsymbol{Y}=(\widetilde{\mathscr{Y}},Y)$ 是定义在 $(\Omega,\mathscr{F},P)$ 上的两个 KF 随机变量。$A\subset R$ 是一个 Borel 集。

定义 4.6.1　在给定 $\boldsymbol{Y}\in A$ 条件下 $\boldsymbol{X}$ 的条件期望 $E(\boldsymbol{X}|\boldsymbol{Y}\in A)$ 为直积 $\boldsymbol{X}\times\boldsymbol{Y}$ 在 R 中的像，映射 $\widetilde{\mathscr{X}}\times\widetilde{\mathscr{Y}}\to R$，由 $(\widetilde{U},\widetilde{V})\to E(\widetilde{U}|\widetilde{V}\in B)$ 确定。

容易证明 $E(\boldsymbol{X}|\boldsymbol{Y}\in B)$ 是一个由下式确定的 F 数：

$$(E(\boldsymbol{X}|\boldsymbol{Y}\in B))(Z)$$
$$=\sup_{\substack{U\in\mathscr{X},\,\pi\in\mathscr{Y}\\E(U\pi)/E(\pi)=z}}\inf_{\omega\in\Omega}\min[\bar{X}_\omega(U(\omega)),\ I_\omega{}^{\boldsymbol{Y}\in B}(\Pi(\omega))],$$

其中 $Z\in R$。

由此推出当 $\boldsymbol{X}$，$\boldsymbol{Y}$ 独立时，$E(\boldsymbol{X}|\boldsymbol{Y}\in A)=E\boldsymbol{X}$。

定义 4.6.2　在 $\boldsymbol{Y}\in B$ 条件下 $\boldsymbol{X}\in A$ 的概率 $P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B)$ 为在映射 $(\widetilde{U},\widetilde{V})\to\widetilde{P}(\widetilde{U}\in A|\widetilde{V}\in B)$ 作用下 $\boldsymbol{X}\times\boldsymbol{Y}$ 于 R 中的像。

由定义，有

$P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B)(p)$

$$= \sup_{\substack{\tilde{U}\in\tilde{\mathscr{X}},\ \tilde{V}\in\tilde{\mathscr{Y}},\\ \tilde{P}(\tilde{U}\in A\mid\tilde{V}\in B)=p}} \inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}} \min[X_{\omega}(\tilde{U}(\omega)), Y_{\omega}(\tilde{V}(\omega))]$$

记 $$\tilde{P}(\tilde{U}\in A|\tilde{V}\in B) = \frac{\tilde{P}(\tilde{U}\in A,\tilde{V}\in B)}{\tilde{P}(\tilde{V}\in B)}$$

$$= \frac{\int dP(\omega)\int_{\tilde{U}\in A,\tilde{V}\in B} dP'(\omega')}{\int dP(\omega)\int_{\tilde{V}\in B} dP'(\omega')}$$

$$= \frac{\int\tilde{\phi}(\omega)\pi(\omega)dP(\omega)}{\int\pi(\omega)dP(\omega)}$$

$$= \hat{E}\tilde{\phi} = \hat{E}\,\hat{E}^{\sigma(\mathbf{X})}\tilde{\phi} = \hat{E}\phi$$

$$= \frac{\int\phi(\omega)\pi(\omega)d\mathscr{P}(\omega)}{\int\pi(\omega)d\mathscr{P}(\omega)},$$

这里

$$\pi(\omega) = \int_{\tilde{V}\in B} d\mathscr{P}'(\omega'),$$

$$\tilde{\phi}(\omega) = \begin{cases} \dfrac{1}{\pi(\omega)}\displaystyle\int_{\tilde{U}\in A,\tilde{V}\in B} d\mathscr{P}'(\omega'), & \Pi(\omega)>0, \\ \text{任意有限} & \Pi(\omega)=0, \end{cases}$$

$$\phi = \tilde{E}^{\sigma(\mathbf{X})}\tilde{\phi},$$

且 $\boldsymbol{X}$ 在映射 $\tilde{U}\to\tilde{\phi}\to\phi$ 之下于 $\mathscr{X}$ 中的像是 $I^{\mathbf{X}\in B}$ 而 $\boldsymbol{Y}$ 在映射 $\tilde{V}\to\Pi$ 之下于 $\mathscr{Y}$ 中的像是 $I^{\mathbf{Y}\in B}$。由上式得

$$P_{r}(\boldsymbol{X}\in A|\boldsymbol{Y}\in B)(p)$$
$$= \sup_{\substack{\Psi\in\mathscr{X},\ \pi\in\mathscr{Y}:\\ E(\Psi\pi)/E\pi=p}} \inf_{\omega\in\Omega} \min[I_{\omega}{}^{\mathbf{X}\in A}(\phi(\omega)), I_{\omega}{}^{\mathbf{Y}\in B}(\pi(\omega))].$$

如同经典概率论叙述的那样，在 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 独立时有

$$P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B)=P_r(\boldsymbol{X}\in A).$$

当 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 为离散的 KF 随机变量时，可给出条件期望与条件概率的计算规则.设 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 由概率集 p_{ij}，$i,j=1,2,\cdots$ 所确定。$\boldsymbol{X}$ 的 F 值由隶属函数 X^i，$i=1,2,\cdots$表示，$\boldsymbol{Y}$ 的由 Y^j，$j=1,2,\cdots$表示。B 为 R 上的 Borel 集。

Ⅰ．确定 $E(\boldsymbol{X}|\boldsymbol{Y}\in B)$的水平集的规则：

步骤 1　对 $j=1,2,\cdots$确定数

$$s_j'=\sup_{y\in B}Y^j(y),\quad s_j''=\sup_{y\in B^C}Y^j(y).$$

步骤 2　确定水平 μ_k，$k=0,1,2,\cdots$，它们与 1 和 $\min(s_k',s_k'')$，$k=1,2,\cdots$不同。整理 μ_k 为顺序

$$1=\mu_0>\mu_1>\mu_2>\cdots.$$

步骤 3　选择一个 $\mu\in[0,1]$.最好一开始取为 μ_k，$k=0,1,\cdots$，则 $\mu=0$，而后取中间值。

步骤 4　确定集合

$$J_\mu'=\{j\in N|s_j'<\mu\},\qquad J_\mu''=\{j\in N|s_j''<\mu\}.$$

$$\bar{J}_\mu'=\{j\in N|s_j'\leqslant\mu\},\qquad \bar{J}_\mu''=\{j\in N|s_j''\leqslant\mu\}.$$

步骤 5　对 $i=1,2,\cdots$确定数

$$u_i^*(\mu)=\inf\{x\in R|X^i(x)\geqslant\mu\},$$

$$u_i^{**}(\mu)=\sup\{x\in R|X^i(x)\geqslant\mu\}.$$

步骤 6　确定数

$$a_\mu=\min\left\{\frac{\sum_{i,j}p_{ij}\pi_j u_i^*(\mu)}{\sum_{i,j}p_{ij}\pi_j}\;(\forall j\in N)\pi_j\in[0,1],\right.$$

$$\pi_j=0,\ \text{若}\ j\in J'_\mu,\ \pi_j=1\ \text{若}\ j\in J''_\mu\Bigg\},$$

$$b_\mu=\max\left\{\frac{\sum_{i,j}p_{ij}\pi_j u_i^{**}(\mu)}{\sum_{i,j}p_{ij}\pi_j}\ \middle|\ (\forall j\in N)\pi\in[0,1],\right.$$

$$\pi_j=0\ \text{若}\ j\in J'_\mu,\ \pi_j=1\ \text{若}\ j\in J''_\mu\Bigg\},$$

$$\alpha_\mu=\min\left\{\frac{\sum_{i,j}p_{ij}\pi_j u_i^{*}(\mu)}{\sum_{i,j}p_{ij}\pi_j}\ \middle|\ (\forall j\in N)\ \pi_j\in[0,1],\right.$$

$$\pi_j=0\ \text{若}\ j\in \bar{J}'_\mu,\ \pi_j=1\ \text{若}\ j\in \bar{J}''_\mu\Bigg\},$$

$$\beta_\mu=\max\left\{\frac{\sum_{i,j}p_{ij}\pi_j u_i^{**}(\mu)}{\sum_{i,j}p_{ij}\pi_j}\ \middle|\ (\forall j\in N)\ \pi_j\in[0,1],\right.$$

$$\pi_j=0\ \text{若}\ j\in \bar{J}'_\mu,\ \pi_j=1\ \text{若}\ j\in \bar{J}''_\mu\Bigg\}.$$

步骤 7 定义

$$C_\mu=\{z\in R\,|\,(E(\boldsymbol{X}|\boldsymbol{Y}\in B))(z)\geqslant\mu\},$$
$$D_\mu=\{z\in R\,|\,(E(\boldsymbol{X}|\boldsymbol{Y}\in B))(z)>\mu\},$$

和集合

$$C_\mu=[a_\mu,b_\mu],\qquad D_\mu=(\alpha_\mu,\beta_\mu).$$

步骤 8 返回至步骤 3

在这个规则的证明之前，先作如下解释：当 $\mu=\mu_k$，$k=1,2,\cdots$，有 $J'_\mu=\bar{J}'_\mu$，$J''_\mu=\bar{J}''_\mu$，所以 $a_\mu=\alpha_\mu$，$b_\mu=\beta_\mu$。

$E(\boldsymbol{X}|\boldsymbol{Y}\in B)$ 的隶属函数是一阶梯形式．$\mu=\mu_k$，C_μ 和 D_μ 的确定足以决定该梯形的范围，除过那些远离梯形的隶属函数的点之外。a_μ，b_μ，α_μ 和 β_μ 的确定，要求对一个特殊类型的分式区间规划问题求解。而 $E(\boldsymbol{X}|\boldsymbol{Y}\in B)$的水平集由足够多的水平数得到，隶属函数可由分解公式(4.4.15)近似地恢复。

证 由上面定理结论知$(E(\boldsymbol{X}|\boldsymbol{Y}\in B))(z)$为

$$\sup\left\{\inf_{i\in N,\, j\in N}\min[\overline{X}^i(u_i),I^j(\pi_j)]\ \middle|\ (\forall i\in N),u_i\in R,\right.$$

$$\left.(\forall j\in N)0\leqslant\pi_j\leqslant1,\ \frac{\sum\limits_{i,j}p_{ij}\pi_j u_i}{\sum\limits_{i,j}p_{ij}\pi_j}=z\right\}.$$

其中 $\overline{X}^i$ 和 X^i 是单峰态的，I 是 $\boldsymbol{Y}\in B$ 的指标函数。因此 $C_\mu=[a_\mu,b_\mu]$，其中

$$a_\mu=\inf\left\{\frac{\sum\limits_{i,j}p_{ij}\pi_j u_i^*(\mu)}{\sum\limits_{i,j}p_{ij}\pi_j}\ \middle|\ (\forall j\in N)0\leqslant\pi_j\leqslant1,\right.$$

$$\left.I^j(\pi_j)\geqslant\mu\right\}$$

$$b_\mu=\sup\left\{\frac{\sum\limits_{i,j}p_{ij}\pi_j u_i^{**}(\mu)}{\sum\limits_{i,j}p_{ij}\pi_j}\ \middle|\ (\forall j\in N)\ 0\leqslant\pi_j\leqslant1,\right.$$

$$\left.I^j(\pi_j)\geqslant\mu\right\}$$

这里 u_i^*，u_i^{**} 见步骤 4。容易验证条件 $I^i(\pi_j)\geqslant\mu$ 与 $\pi_j=0$ $(j\in J'_\mu)$；$\pi_j=1(j\in J''_\mu)$；$0\leqslant\pi_j\leqslant1$(其它)等价。由此而得步骤 6 中之 a_μ 及 b_μ。D_μ 的情形类似可证。

例 4.6.1 在关于某地气候的意见征集问题中。我们再考虑一个问题，问旅游者在此期间是否可很好的度假。可能的回答是“好”，“一般”及“没意见”。其隶属函数表示这些答复是确定于 0（绝对不满意）至 1（完全满意）这个尺速范围之间。如图 4.6.1 所示。

表 2　　p_{ij} 值

	好假期 Y^1	一　般 Y^2	没意见 Y^3
很　暖 X^1	0.3	0.1	0
暖 X^2	0.2	0.25	0.5
不评价 X^3	0	0.05	0.05

表 2 就 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 定义了联合 F 随机变量。其中 $\boldsymbol{X}$ 是关于气候，$\boldsymbol{Y}$ 是关于季节问题。现就 $B=[0.8,1]$来计算 $E(\boldsymbol{X}|\boldsymbol{Y}\in B)$。

1. $S'_1=1$，　$S'_2=0$，　$S'_3=1$；
   $S''_1=0.5$，　$S''_2=1$，　$S''_3=1$。
2. $\min(S'_k,S''_k)=0$，故 $\mu_0=1>0.5>0$。
3. 取 $\mu=\mu_0=1$。
4. $J'_1=\{2\}$，$J''_1=\{1\}$；$\bar{J}'_1=\{1,2,3\}$. $\bar{J}''_1=\{1,2,3\}$。
5. $u_1^*(1)=30$，　$u_1^{**}(1)=30$，

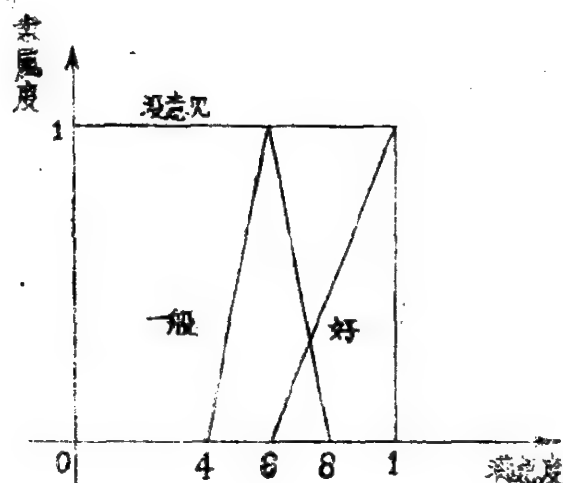


图 4.6.1

$$u_2^*(1)=25, \qquad u_2^{**}(1)=25,$$

$$u_3^*(1)=25, \qquad u_3^{**}(1)=25.$$

6. $a_1=27.5$，$b_1=28$ 从而 $C_1=[27.5,\ 28]$，类似可得 $D_{0.5}=(25,\ 30.5)$，$C_{0.5}=[22.5,\ 30.5]$和$D_0=(20,\ 33)$. 对于 μ 在 0.5 与 1 及 0 与 0.5 之间的值亦可确定 C_μ.从而完成$E(\boldsymbol{X}\,|\,\boldsymbol{Y}\in B)$的图像.

如果取 $B=[0,\ 1]$，可得 $C_1=[25,\ 28]$，$D_0=(20,\ 32.5)$，$C_{0.5}=[22.5,\ 30.5]$，$D_{0.5}=(22.5,\ 30.5)$。从而得 $E(\boldsymbol{X}\,|\,\boldsymbol{Y}\in B)$之略图.它与上图有明显的差别. 此时 $E(\boldsymbol{X}\,|\,\boldsymbol{Y}\in B)$即每个满意的程速，在其上个别的满意与他的假期为条件是等价的，并且是完全可以接受的.

注 步骤 6 的计算用到分式区间规划问题的求解方法，这里从略.

Ⅱ. *F* 条件概率的计算规则

设 A 和 B 是 R 中的 Borel 集，KF 随机变量 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$的

联合概率分布为 p_{ij}，$i\in N$，$j\in N$，且其隶属函数分别为 X^i，$i\in N$ 和 Y^j，$j\in N$。

步骤 1 对每个 i 和 j 确定数

$$r_i' = \sup_{x\in A} X^i(x), \qquad r_i'' = \sup_{x\in A^c} X^i(x),$$

$$s_j' = \sup_{y\in B} Y^j(y), \qquad s_j'' = \sup_{y\in B^c} Y^j(y).$$

步骤 2 确定水平 μ_k，$k=0, 1, 2, \cdots$。设 $(P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B))$ 与 1、$\min(r_i',r_i'')$ 和 $\min(s_j',s_j'')$ 是不同的值.整理 μ_k 成 $1=\mu_0>\mu_1>\mu_2>\cdots$。

步骤 3 对 $k=0, 1, 2, \cdots$ 确定整数集

$$I_k' = \{i\in N | r_i' < \mu_k\}, \quad I_k'' = \{i\in N | r_i'' < \mu_k\},$$

$$J_k' = \{j\in N | s_j' < \mu_k\}, \quad J_k'' = \{j\in N | s_j'' < \mu_k\}.$$

步骤 4 对于 $k=0, 1, \cdots$ 确定函数

$$\chi_k(q) = \begin{cases} 1, & q\in[q_k^*, q_k^{**}], \\ 0, & \text{其它}. \end{cases}$$

这里

$$q_k^* = \min\left\{ \frac{\sum\limits_{j\in J_k'\cup J_k''} \pi_j \sum\limits_{i\in I_k''} p_{ij} + \sum\limits_{j\in J_k''} \sum\limits_{i\in I_k''} p_{ij}}{\sum\limits_{j\in J_k'\cup J_k''} \pi_j \sum\limits_{i} p_{ij} + \sum\limits_{j\in J_k''} \sum\limits_{i} p_{ij}} \,\middle|\, (\forall j\in N)\pi_j\in[0,1] \right\},$$

$$q_k^{**} = \max\left\{ \frac{\sum\limits_{j\in J_k'\cup J_k''} \pi_j \sum\limits_{i\notin I_k'} p_{ij} + \sum\limits_{j\in J_k''} \sum\limits_{i\notin I_k'} p_{ij}}{\sum\limits_{j\in J_k'\cup J_k''} \pi_j \sum\limits_{i} p_{ij} + \sum\limits_{j\in J_k''} \sum\limits_{i} p_{ij}} \,\middle|\right.$$

$$(\forall j\in N)\pi_j\in[0,1]\Big\}.$$

步骤 5 最后，对 $q\in[0,1]$

$$(P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B))(q)=\sup_k \chi_k(q)\mu_k.$$

证 由定义 4.6.2 的结论，对 $q\in[0,1]$有

$$(P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B))(q)$$

$$=\sup\Bigg\{\inf_{i\in N,j\in N}\min[I^i(\phi_i),J^j(\pi_j)]\Bigg|$$

$$(\forall i\in N,\forall j\in N)\ 0\leqslant\phi_i\leqslant 1,0\leqslant\pi_j\leqslant 1,$$

$$\frac{\sum_{i,j}p_{ij}\phi_i\pi_j}{\sum_{i,j}p_{ij}\pi_j}=q\Bigg\}$$

这里，I^i，$i\in N$是指标函数$I^{X\in A}$的隶属函数，且Y^j，$j\in N$是$I^{Y\in B}$的。现在我们来确定水平集$C_\mu=\{q\in[0,1]|(P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B))(q)\geqslant\mu\}$。则对于$\mu\in[0,1]$，有$C_\mu=[a_\mu,b_\mu]$，这里

$$a_\mu=\inf\Bigg\{\frac{\sum_{i,j}p_{ij}\pi_j\phi_i^*(\mu)}{\sum_{i,j}p_{ij}\pi_j}\Bigg|(\forall j\in N)\ 0\leqslant\pi_j\leqslant 1,$$

$$J^j(\pi_j)\geqslant\mu\Bigg\},$$

$$b_\mu=\sup\Bigg\{\frac{\sum_{i,j}p_{ij}\pi_j\phi_i^{**}(\mu)}{\sum_{i,j}p_{ij}\pi_j}\Bigg|(\forall j\in N)\ 0\leqslant\pi_j\leqslant 1,$$

$$J^i(\pi_j)\geqslant\mu \Big\},$$

其中

$$\phi_i^*(\mu)=\inf\{\phi\in[0,1]\,|\,I^i(\phi)\geqslant\mu\},$$

$$\phi_i^{**}(\mu)=\sup\{\phi\in[0,1]\,|\,I^i(\phi)\geqslant\mu\}.$$

借助 $I^i(\pi)$容易求得

$$\phi_i^*(\mu)=\begin{cases}0, & r_i''\geqslant\mu,\\ 1, & r_i''<\mu,\end{cases}$$

$$\phi_i^{**}(\mu)=\begin{cases}0, & r_i'<\mu,\\ 1, & r_i'\geqslant\mu.\end{cases}$$

此外条件 $J^i(\pi_j)\geqslant\mu$ 包括

$$\phi_j=\begin{cases}0, & s_j'<\mu,\\ 1, & s_j''<\mu,\\ \text{任意}, & \text{其它}.\end{cases}$$

从作为 μ 的函数的表达式得出水平集 C_μ 仅仅在由规则步骤 2 中指出的水平 μ_k，$k=0$，1，2，…处发生改变. 确定 $a_{\mu_k}=q_k^*$，$b_{\mu_k}=q_k^{**}$，采用在步骤 3 中引入的整数集而直接得到步骤 4 中给出的两个表达式. 末了从一般表达式 $m(\alpha)$给出步骤 5 中关于$(P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B))$的表达式.

注 步骤 5 中 q_k^* 和 q_k^{**} 的确定要求一个分段区间规划问题的解.

例 4.6.2 考虑例 1 中的联合分布 F 随机变量，计算 F 条件概率 $P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B)$，其中 $A=[z,\infty)$且$B=[0.8,1]$. 如此，我们计算个别答复温度超过 z 的 F 条件概率，条件为他对假期满意是在 0.8 与 1 之间. 此外应用上面给定的规则. 对于 $z=27$℃，例如，求得 $\mu_0=1$，$\mu_1=0.6$，$\mu_2=0.5$，

$\mu_4=0$. 整数集 $I'_0=\{2\}, I''_0=\{1\}, J'_0=\{2\}, J''_0=\{1\}$, $I'_1=\phi$, $I''_1=\{1\}$, $J'_1=\{2\}$, $J''_1=\{1\}$, $I'_2=\phi$, $I''_2=\{1\}$, $J'_2=\{2\}$, $J''_2=\phi$, $I'_3=\phi$, $I''_3=\phi$, $J'_3=\{2\}$, $J''_3=\phi$, $I'_4=I''_4=J'_4=J''_4=\phi$. 由此我们得 $c_1=[0.5,\ 0.6]$, $c_{0.6}=[0.5,\ 1]$, $c_{0.5}=c_{0.4}=c_0=[0,1]$。

对于不同的 z 值($z=26,28,29,30$等)类似上面步骤可算出在条件 $\boldsymbol{Y}\in B$ 下的 F 条件概率，将其与没有此条件的 F 概率 $P_r(\boldsymbol{X}\in A)$ 对应比较，将发现条件影响不大(参见[20])。

当 A 和 B 自身是不分明时，我们可以推广 F 条件概率的概念。令 $\boldsymbol{A}$ 是定义于 R 中 Borel 集的空间$\mathscr{B}$上的不分明集，$A\in\mathscr{B}$在$\boldsymbol{A}$中的隶属度由 $\alpha(A)$ 给出，这里 α: $\mathscr{B}\to[0,1]$。类似，令 $\boldsymbol{B}$表示定义在$\mathscr{B}$上，使得 $B\in\boldsymbol{B}$ 在$\boldsymbol{B}$中具有隶属度 $\beta(B)$ 的 F 集，其中 β: $\mathscr{B}\to[0,1]$。那么，我们定义 $P_r(\boldsymbol{X}\in \boldsymbol{A}|\boldsymbol{Y}\in\boldsymbol{B})$为一个具有隶属函数

$$(P_r(\boldsymbol{X}\in\boldsymbol{A}|\boldsymbol{Y}\in\boldsymbol{B}))(\cdot)=\sup_{A,B\in\mathscr{B}}\min[(P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B))(\cdot),\ \alpha(A),\ \beta(B)]$$

的 F 数。

Ⅲ. $P_r(\boldsymbol{X}\in A|\boldsymbol{Y}\in B)$的计算规则

步骤1 对每个 $i\in N$ 和 $j\in N$ 确定数值

$$r'_i=\sup_{A\in\mathscr{B}}\min[\alpha(A),\ \sup_{x\in A}X^i(x)],$$

$$r''_i=\sup_{A\in\mathscr{B}}\min[\alpha(A),\sup_{x\in A^c}X^i(x)],$$

$$s'_j=\sup_{B\in\mathscr{B}}\min[\beta(B),\sup_{y\in B}Y^j(y)],$$

$$s''_j=\sup_{B\in\mathscr{B}}\min[\beta(B),\sup_{y\in B^c}Y^j(y)].$$

步骤 2 至步骤 4　如同规则Ⅱ中的步骤 2 至步骤 4那样。

步骤 5　末了，对 $q\in[0,1]$确定

$$(P_r(\boldsymbol{X}\in\boldsymbol{A}|\boldsymbol{Y}\in\boldsymbol{B}))(q)=\sup_k X_k(q)\mu_k.$$

如此，规则Ⅲ与规则Ⅱ实际是相同的，除过步骤 1 中r_i'，r_i''，s_i'和s_i''改变了之外。规则Ⅲ的证明是不困难的。

§7　应用举例

4.7.1　在简单决策问题上的应用

在这一部分，我们讨论对于一个简单决策问题的假想的应用，它是关于医疗诊断的问题。假设一个确定的医学检验是用于决定是否一个人患一种确定的病。按大多数情形，关联到检验疾病存在的结果的统计信息是可以接受的。

由于检验结果分析困难，以及不能用一个单个的数来表示，从而产生不精确性。我们认为检验的结果，理解为阳性、不确定的和阴性，它们被隶属函数表示成为由 0 到 1 范围上的数量。在有不分明情况下隶属函数部分重叠是可能的。在更切合实际的应用中，假定一个比较大的可能分类（严格阳性，适当的阳性，稍有点阳性，严格阴性，等等）将被允许。

类似地，在扩充的指标中有不精确性，这是由于把一个病人的病简化了。我们假定以下三个指标——严重疾病、轻微疾病和没有病，能用隶属函数 σ 表示，如图 4.7.1 所示。

对于疾病检验结果的统计数据汇聚于表 3 之中。正如表中指出，疾病的程度用 F 随机变量 $\boldsymbol{X}$ 来表示，检验的结果用 F 随机变量 $\boldsymbol{Y}$ 来表示。

表 3　　检验和疾病性两者结果间的联系

	检验结果		
	Y^1 否定	Y^2 不确定	Y^3 正确
X^1 无病	0.78	0.07	0.05
X^2 轻病	0.02	0.02	0.01
X^3 严重病	0	0.01	0.04

现在假定对于一个确定的病人检验结果，诊断为不能确实否定。这个意思是指检验结论的依据的平均强度属于一个不分明Bore集合$\boldsymbol{B}$中，$\boldsymbol{B}$的隶属函数为

$$\beta(\boldsymbol{B})=\begin{cases}\sigma(a), & \boldsymbol{B}=[0,a],\\ 0, & \text{其它}.\end{cases}$$

其图像示于图 4.7.1 中.

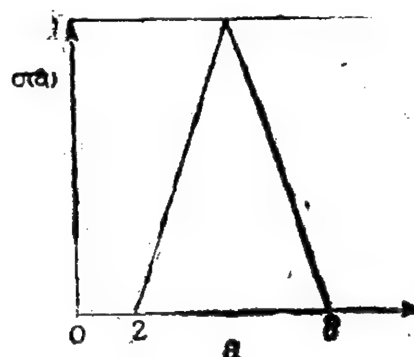


图 4.7.1　隶属函数 σ

现在我们来计算 $P_r(\boldsymbol{X}\in A_i|\boldsymbol{Y}\in\boldsymbol{B})$, $i=1,2,3$, 且

$$A_1[0,\ 0.3],\ A_2=[0.3,\ 0.7],\ A_3=[0.7,\ 1].$$

这里事件 $\boldsymbol{X}\in A_1$ 为“无严重疾病”，$\boldsymbol{X}\in A_2$为“轻病”，而$\boldsymbol{X}\in A_3$为“严重疾病”. 这三个隶属函数可借助规则Ⅲ得到. 它将指出在患者的检验结果条件下无严重疾病的条件概率，似乎有充分理由处于 0.944 和 0.975 之间；其它值似乎真的是$\frac{1}{3}$或小于$\frac{1}{3}$. 另一方面，个别人在给出检验结果条件下为严重疾病的条件概率，似乎真的处于 0.0111 与 0.0556 之间，而其它概率似乎真的为 1/3 或小于 1/3. 这些结果指出个别人

在给定检验结果的条件下无病的概率很大。

注意观察该例中的检验结果是很有意思的一件事情."相对于非确定性为低"，这在汇总统计("否定","不确定"及"正确"的统计)中不是一个检验的结果。然而,我们可以得到(含混的)结论。有几分"否定"和有几分"不确定"的检验结论可由不分明逻辑法则得到。

4.7.2 另一个应用(P.Smets,1982)

假设 z 是定义于 Ω 上的一个随机变量，具有概率测度 $P(z)$。又设 $X=v(z)$, $Y=w(z)$是分别定义于 Ω_1 和 Ω_2上的两个随机变量，具有联合密度函数

$$f_{X,Y}(x,y)=\int_{Z(x,y)} dP(Z),$$

其中

$$Z(x,y)=\{z: z\in\Omega,\ v(z)=x,\ w(z)=y\},$$

并且边际密度函数为 $f_X(x)$, $f_Y(y)$。

设 $\underset{\sim}{A}$ 是一个定义在 Ω 上的不分明集，其隶属函数仅依赖于 $\nu(z)$，即 $\underset{\sim}{A}(z)=h_{\underset{\sim}{A}}(\nu(z))$，则

$$\tilde{P}(\underset{\sim}{A})=\int_{\Omega} h_{\underset{\sim}{A}}(x) f_X(x)dx \qquad (4.7.1)$$

并且

$$\tilde{P}(\underset{\sim}{A}|Y=y)=\int_{\Omega} h_{\underset{\sim}{A}}(x) f_{X,Y}(x,y)dx/f_Y(y)$$

按照 Bayes 定理

$$f_Y(y|\underset{\sim}{A})=\tilde{P}(\underset{\sim}{A}|Y=y)f_Y(y)/\tilde{P}(\underset{\sim}{A})$$

$$=\int_{\Omega_1} h_{\underset{\sim}{A}}(x) f_{X,Y}(x,y)dx/\tilde{P}(\underset{\sim}{A}).$$

在(4.7.1)式中的密度函数 $f_X(x)$ 考虑为集中在 x 的

Dirac 函数，那末 $P(\underset{\sim}{A}|X=x)=h_{\underset{\sim}{A}}(x)$。并且由 Bayes 定理

$$f_X(x|\underset{\sim}{A})=\tilde{P}(\underset{\sim}{A}|x)f_X(x)/\tilde{P}(\underset{\sim}{A})$$
$$=h_{\underset{\sim}{A}}(x)f_X(x)/\tilde{P}(\underset{\sim}{A}).$$

作为一个例子，考虑人类 Z 的集Ω. 设 $v(z)$是 Z 的年令，$w(z)$是 Z 的身高，$\underset{\sim}{A}$ 是老人的不分明集. $f_Y(y|\underset{\sim}{A})$和$f_X(x|\underset{\sim}{A})$是老人的身高和年令的分布密度，而

$$E(Y|\underset{\sim}{A})=\int_{\Omega_2}yf_Y(y|\underset{\sim}{A})dy$$

及

$$E(X|\underset{\sim}{A})=\int_{\Omega_1}xf_X(x|\underset{\sim}{A})dx$$

分别表示老人的平均身高与平均年龄。

考虑两个随机变量 X 和 Y。定义在 Ω_1 和 Ω_2 上，具有联合密度 $f_{X,Y}(x,y)$。假设在 $\Omega_1\times\Omega_2$ 上一个不分明关系R，其隶属函数是 $\mu_R(x,y)$，$x\in\Omega_1$，$y\in\Omega_2$。定义

$$\tilde{P}(R)=P\{(x,y)\in R\}$$
$$=\int_{\Omega_1}\int_{\Omega_2}\mu_R(x,y)f_{X,Y}(x,y)dydx$$

且$\tilde{P}(R|x,y)$是当$f_{X,Y}(x,y)$是一个集中于 x,y的二维Dirac 函数时，得到$\tilde{P}(R|x,y)=\mu_R(x,y)$。

由 Bayes 定理，给出不分明关系 R 的 X 和 Y 的条件联合分布是

$$f_{X,Y}(x,y|R)=\tilde{P}(R|x,y)f_{X,Y}(x,y)/\tilde{P}(R)$$
$$=\mu_R(x,y)f_{X,Y}(x,y)/\tilde{P}(R).$$

条件密度函数

$$f_X(x|y,R)=\mu_R(x,y)f_{X,Y}(x,y)\Big/\Big(\int_{\Omega_1}\mu_R(x,y)$$

$$\cdot f_{X,Y}(x,y)dx\Big)$$

$$=\mu_R(x,y)f_{X|Y}(x,y)\Big/\Big(\int_{\Omega_1}\mu_R(x,y)\cdot f_{X|Y}(x|y)dx\Big).$$

作为一个例子，考虑人的身高X和体重Y，设它们服从均值为μ，方差矩阵为$\sum$的二维正态分布。X单位用厘米表示、Y的单位用公斤表示。考虑关系R，其X和Y的数值相互靠近，即$|x-y|$小，$\mu_R(x,y)=\exp[-(x-y)^2/50]$。特别，$\mu_R(x,x)=1$，$\mu_R(x,x+10)=0.14$，$R$暗示$|X-Y|$实质上比10小。令

$$V^{-1}=0.04\begin{bmatrix}1, & -1\\ -1, & 1\end{bmatrix},$$

$$A^{-1}=V^{-1}+\sum{}^{-1}.$$

那么$f_{X,Y}(x,y|R)$是具有均值$A\sum^{-1}\mu$和协方差矩阵A的双变数正态。假如

$$\mu'=(70,70),\quad \sum=\begin{bmatrix}100, & 60\\ 60, & 225\end{bmatrix}$$

则

$$A\sum{}^{-1}\mu=\mu,\qquad A=\begin{bmatrix}93, & 89\\ 89, & 107\end{bmatrix},$$

R忽略稳定的均值，省去两者的方差，其相关程度从0.40增加到0.89。

第五章

语言概率

研究语言概率的目的，是为了将人类的语言用于自动机器，更好的发挥人的作用. 这个课题仍然是 Zadeh 首先引进的，后来得到不断的发展，并在人机系统中得到应用. 从目前发展的情况看来，离散情况刻划得比较好，而对于连续情况，尚有进一步研究的必要. 研究语言概率的理论基础，仍然是概率论与不分明集理论. 语言概率也成了一种特殊的不分明集.

人类的语言，基本上是模糊的，但是却能在人们的实践活动中完成其目的. 另外，在系统工程中，经常考虑的对象是人. 为了更充分的发挥人的作用，模拟人的思维能力，就必须研究人的直觉和判断，研究人类的语言结构. 例如，人们在日常生活中，经常遇到这样一个问话：甲乙两队进行比赛，问甲队赢的可能性多大？此种情况往往得到的回答是，甲队赢的可能性"很大"或"很小". 仔细想来，这里的"很大"与"很小"的概念，不是一个精确的概念，而是不分明的. 但是，"甲队赢"是一个明确的概念，"甲队赢的可能性很大"却涉及一个确切的事件的"不确切"概率的问题.

在日常生活中，像这样的语言很多，如"根本不可能"，"不很可能"，"很不可能"，"很可能"，"非常可能"，"完全可能"等等都是如此. 当我们说"根本不可能"时，意思是"该事件出现的概率几乎就是零"，而"很不可能"指"出现的概率

很小”，“非常可能”却指“该事件出现的概率很大”等等。

本章先介绍不分明语言及其语义结构、语言变量，最后介绍语言概率。主要取材于 Zadeh、汪培庄及Lee 的文章。这些对进一步理解前几章的内容将是有益的。

§1 不分明语言

我们知道语言是由句子构成，而句子则是由词经过一定的语法规则来构成。因而单词就形成了语言表达中的最小单位。另一方面，词又是由字母所组成，从而使得句子不外是字母的一个有限序列。而不分明语言就成了字母的有限序列的集合上的不分明子集。

下面先考虑不分明语言的语义结构。

5.1.1 单词

单词是人类语言中表达概念的最小单位，如人、尺、手、快、慢、强、弱、黑、白、大、小，等等。它们相对于一定的范畴，这个范畴我们称之为论域，记作U。它是词义的集。而一类单词构成一个集合 T。因此，单词的语义（即单词的含义）应该为 T 到 U 上的一个对应关系。

例如，要在论域 $U=\{1,2,3,4,5\}$上谈论“大”“小”这两个词，那么，可以定义

$$[\text{大}]=0.5/4+1/5,$$

$$[\text{小}]=1/1+0.2/2.$$

可见 $T=\{[\text{大}],[\text{小}]\}$，而[大] 的语义却是从 T 至 U 的一个映射。

由此可以得出单词的语义是一个映射：

$$N(a,u): T\rightarrow U,$$

$\alpha \in T$，$u \in U$。从而 $N(\alpha, u)$ 是 U 上的一个不分明集 $\underset{\sim}{A}(u)$。仅当 $\underset{\sim}{A} = A$ 时，称单词 α 为确定的。

5.1.2 词组

单词通过逻辑运算可以构成新词，这个新词称为词组。

单词可看作是论域上的不分明集，从而词组可以看作是不分明集逻辑运算的结果。这里所指的逻辑运算一般是指"∩"、"∪"及"—(非)"的运算。例如：

$$[白马] = [白] \cap [马];$$
$$[东南方] = [东方] \cap [南方];$$
$$[非金属] = \overline{[金属]}.$$

5.1.3 算子

在自然语言中有些词，如"很"、"稍许"、"极"、"略"、"非常"、"微"、"特别"等等，如果将其缀在一个单词(譬如说"老"这个单词）的前面，便调整了该词的词义获得新词：如"很老"、"极老"，"比较老"，"特别老"等等。这类字眼，我们可以把他看作是一种算子。

常用的算子有以下几种

(1) 语气算子

有那么一些词，它们作用于另一单词后，使得该单词的语义发生了语气上的变化，这类词我们称它为语气算子。象"很"、"极"、"微"、"略"、"稍许"等词，都是语气算子。一旦将它们作用于单词"大"，那么得到的新词与"大"有关，但大的程度发生了变化。我们不妨取"很"为 H_2，那么

$$\begin{aligned}[很大](u) &= [很]([大](u)) \\ &= H_2([大](u))\end{aligned}$$

令

$$H_2([大](u))=([大](u))^2.$$

仿此可定义其它情形.

从而我们定义，一个语气算子是一个映射H_λ: $\mathscr{F}(U)\to\mathscr{F}(U)$。其中 λ 为正数,而$\mathscr{F}(U)$由论域 U 的子集所组成。一般取形式

$$(H_\lambda \underset{\sim}{A})(u)=[\underset{\sim}{A}(u)]^\lambda.$$

并且，当 $\lambda>1$ 时，H_λ 称为集中化算子；当 $\lambda<1$ 时，称 H_λ 为散漫化算子。

对于不同的 λ 值，我们可以规定它们的名字。如,H_2 叫做"很"，"$H1/2$"叫做"略"，"H"叫做"极"，"$H1/4$"叫做"微"等等。

例 5.1.1 令[老]:

$$[老](u)=\begin{cases}0, & (0<u<50)\\ \left[1+\left(\dfrac{u-50}{5}\right)^2\right]^{-1}, & (50<u<100)\end{cases}$$

于是

$$[很老](u)=H_2[老](u)=([老](u))^2$$

$$=\begin{cases}0 & (0<u<50)\\ \left\{1+\left(\dfrac{u-50}{5}\right)^2\right\}^{-2} & (50<u<100).\end{cases}$$

$$[极老](u)=H_4[老](u)=([老](u))^4$$

$$=\begin{cases}0 & (0<u<50)\\ \left(1+\left(\dfrac{u-50}{5}\right)^2\right)^{-4} & (50<u<100).\end{cases}$$

$$[略老](u)=H_{1/2}[老](u)=([老](u))^{1/2}$$

$$=\begin{cases}0 & (0<u<50)\\ \left(1+\left(\frac{u-50}{5}\right)^2\right)^{-1/2} & (50<u<100)\end{cases}$$

$$\begin{aligned}[\text{很很老}](u)&=H_2[\text{很老}](u)=H_2[H_2[\text{老}]](u)\\&=[([\text{老}](u))^2]^2=([\text{老}](u))^4\\&=[\text{极老}](u).\end{aligned}$$

等等。

(2) 模糊化算子

一个词经过算子作用以后，该词的语义变得模糊，这类算子称为模糊化算子。如“几乎”、“大概”、“近乎”等词。用 F 表示模糊化算子。

模糊化算子的一般形式是

$$\begin{aligned}(F\underset{\sim}{A})(u)&=(F\circ\underset{\sim}{A})(u)\\&=\bigvee_{v\in U}(\underset{\sim}{E}(u,v)\wedge\underset{\sim}{A}(u))\end{aligned}$$

此处$\underset{\sim}{E}$是 U 上的一个相似关系，当 $U=R=(-\infty,+\infty)$时，常取

$$\underset{\sim}{E}(v,u)=\begin{cases}e^{-(u-v)^2}, & |u-v|<\delta\\ 0, & |u-v|>\delta.\end{cases}$$

(δ 是参数)。

例 5.1.2 令

$$\underset{\sim}{A}(x)=\begin{cases}1, & x=3.0,\\ 0, & x\neq3.0.\end{cases}$$

即 $\underset{\sim}{A}$ 是“3”字的“义”，则

$$F\underset{\sim}{A}(x)=\bigvee_{v\in X}(\underset{\sim}{E}(x,y)\wedge\underset{\sim}{A}(x))$$

$$=\underset{\sim}{E}(x,3.0)$$

$$=\begin{cases} e^{-(x-3.0)^2} & |x-3.0|\leqslant\delta, \\ 0 & |x-3.0|>\delta. \end{cases}$$

$F\underset{\sim}{A}$对应的词叫做“大约 3”。如图 5.1.1 所示。

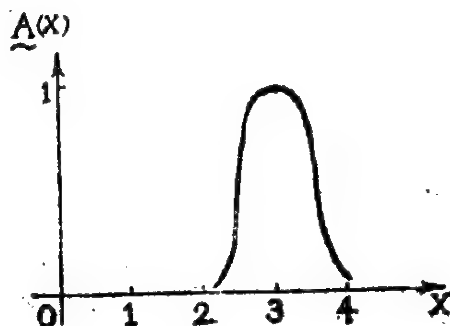


图 5.1.1 $\delta=0.5$ 的[大约 3]

（3） 判定化算子

将一个不分明词转化为肯定的算子，称为判定化算子。这是由于它给出在不分明之中一个粗糙的判断。如象“偏向”、“倾向于”、“多半是”等词。

一般地用 P_a 表示判定化算子。定义为

$$P_a\underset{\sim}{A}(u)=d_a[\underset{\sim}{A}(u)]$$

此处 d_a 是[0,1]上的实函数：

$$d_a(x)=\begin{cases} 0, & x\leqslant a, \\ \dfrac{1}{2}, & a<x\leqslant 1-a, \\ 0, & x>1-a. \end{cases}$$

而 a 为$\left(0,\dfrac{1}{2}\right]$之间的一个数。

当 $a=\dfrac{1}{2}$时的 $P_{1/2}$叫做“偏向”算子。例如

$$[\text{偏向年轻}](u)=(P_{1/2}[\text{年轻}])(u)$$
$$=d_{\frac{1}{2}}([\text{年轻}](u)).$$

如果认为

$$[年轻](35)=\frac{1}{2}。$$

则

$$[偏向年轻](u)=\begin{cases}0, & u>35,\\ 1, & u\leqslant 35。\end{cases}$$

当 $a<\frac{1}{2}$ 时，这种判断是留有余地的。当隶属度介于 a 与 $1-a$ 之间时，反而不分明化为 $\frac{1}{2}$。

现在我们把不分明语言看作是字母的有限序列的集合上的不分明子集。依此来讨论不分明语言的造句（Syntactical）。

设集合 $X\neq\phi$，为字母(letter)的集合，L 为具有 0 和 1 的分配格。

定义 5.1.1 设由字母的序列构成的集合为 X^*，X上的不分明语言是 X^*上的不分明子集。而 $\mathscr{L}: X^*\to L$。

我们将所有不分明语言的集合表示为 $\mathscr{L}(X)=\mathscr{F}(X^*)$。

在集合 $\mathscr{L}(X)$ 上定义"和"运算"$\vee$"与"积"运算"$\wedge$"为：

$$(\mathscr{L}\vee\mathscr{L}')(\theta)=\mathscr{L}(\theta)\vee\mathscr{L}'(\theta)$$

$$(\mathscr{L}\wedge\mathscr{L}')(\theta)=\mathscr{L}(\theta)\wedge\mathscr{L}'(\theta)$$

其中 $\mathscr{L}$，$\mathscr{L}'\in\mathscr{L}(X)$，$\theta\in X$。

显然 $\mathscr{L}(X)$ 对上述定义的运算"$\wedge$""$\vee$"构成一分配格。

另一个重要的运算是不分明语言的串连(concatenation)定义为

$$\mathscr{L},\mathscr{L}'\in\mathscr{L}(X)\Rightarrow\mathscr{L},\mathscr{L}'\in\mathscr{L}(X),$$

$$(\mathscr{L}\mathscr{L}')(\theta)=\bigvee_{\substack{\theta',\theta''\in X^*\\ \theta'\theta''=\theta}}[\mathscr{L}(\theta')\wedge\mathscr{L}'(\theta'')],\ \theta\in X^*$$

定理 5.1.1 不分明语言的串联满足结合律。

证 设$\mathscr{L},\mathscr{L}',\mathscr{L}''\in\mathscr{L}(X)$，我们要证明有

$$(\mathscr{L}\mathscr{L}')\mathscr{L}''=\mathscr{L}(\mathscr{L}'\mathscr{L}'')。$$

因为

$$[(\mathscr{L}\mathscr{L}')\mathscr{L}''](\theta)=\bigvee_{\theta'\theta_3=\theta}[(\mathscr{L}\mathscr{L}')(\theta')\wedge\mathscr{L}''(\theta_3)]$$

$$=\bigvee_{\theta'\theta_3=\theta}\left\{\left[\bigvee_{\theta_1\theta_2=\theta'}(\mathscr{L}(\theta_1)\wedge\mathscr{L}'(\theta_2))\right]\wedge\mathscr{L}''(\theta_3)\right\}$$

$$=\bigvee_{\theta_1\theta_2\theta_3=\theta}[\mathscr{L}(\theta_1)\wedge\mathscr{L}'(\theta_2)\wedge\mathscr{L}''(\theta_3)]$$

而$\mathscr{L}(\mathscr{L}'\mathscr{L}'')$也得到同样的表示式。从而结合律得证。

定义 5.1.2 空语言(empty language)$\mathscr{L}_\Lambda\in\mathscr{L}(X)$为

$$\mathscr{L}_\Lambda(\theta)=\begin{cases}1, & \theta=\Lambda,\\0, & \theta\neq\Lambda.\end{cases}$$

当L为完备时，我们可以定义不分明语言的Kleene闭包(closure)。

定义 5.1.3 不分明语言$\mathscr{L}\in\mathscr{L}(X)$的Kleene闭包是不分明语言$\widehat{\mathscr{L}}$：

$$\widehat{\mathscr{L}}=\mathscr{L}_\Lambda\vee\mathscr{L}\vee\mathscr{L}\mathscr{L}\vee\cdots=\bigvee_{j=0}^{\infty}\mathscr{L}^j。$$

注 一般，若不分明语言$\mathscr{L}\in\mathscr{L}(X)$为独异点同态(monoid homomorphism，或单半群同态)，即

$$\mathscr{L}(\Lambda)=1$$

$$\mathscr{L}(\theta\theta')=\mathscr{L}(\theta)\wedge\mathscr{L}(\theta')$$

时，有$\widehat{\mathscr{L}}=\mathscr{L}$，(即是“闭语言”)。这可由

$$(\mathscr{L}\mathscr{L})(\theta)=\bigvee_{\theta'\theta''=0}[\mathscr{L}(\theta')\wedge\mathscr{L}(\theta'')]$$

$$=\bigvee_{\theta'\theta''=\theta}[\mathscr{L}(\theta'\theta'')]$$

$$=\mathscr{L}(\theta)$$

及$\mathscr{L}_\Lambda\leqslant\mathscr{L}$得

$$\hat{\mathscr{L}}=\mathscr{L}_\Lambda\vee\mathscr{L}\vee\mathscr{L}\mathscr{L}\vee\cdots$$

$$=\mathscr{L}_\Lambda\vee\mathscr{L}=\mathscr{L},$$

从而得到结论。

我们可给出一个更确切的结论：闭语言的特征定理。

定理 5.1.2 设$\mathscr{L}\in\mathscr{L}(X)$是不分明语言，如下两个命题是等价的。

(i) $\hat{\mathscr{L}}=\mathscr{L}$.

(ii) $\mathscr{L}$ 满足条件 $\mathscr{L}(\Lambda)=1$,

$$\mathscr{L}(\theta\theta')\geqslant\mathscr{L}(\theta)\wedge\mathscr{L}(\theta').$$

证 首先证明(i)⇒(ii)。

根据所设$\hat{\mathscr{L}}=\mathscr{L}$，则$\mathscr{L}_\Lambda\leqslant\mathscr{L}$，因此$\mathscr{L}(\Lambda)=1$。又因为

$$(\mathscr{L}\mathscr{L})(\theta)=\bigvee_{uv=\theta}[\mathscr{L}(u)\wedge\mathscr{L}(v)]$$

$$\geqslant\mathscr{L}(\theta)\wedge\mathscr{L}(\Lambda)=\mathscr{L}(\theta),$$

因此，$\mathscr{L}\mathscr{L}\geqslant\mathscr{L}=\hat{\mathscr{L}}\geqslant\mathscr{L}\mathscr{L}$，而有$\mathscr{L}=\mathscr{L}\mathscr{L}$，即

$$\mathscr{L}(\theta)=\bigvee_{uv=\theta}[\mathscr{L}(u)\wedge\mathscr{L}(v)]$$

成立，于是

$$\mathscr{L}(\theta\theta')=\bigvee_{uv=\theta\theta'}[\mathscr{L}(u)\wedge\mathscr{L}(v)]$$

$$\geqslant \mathscr{L}(\theta)\wedge\mathscr{L}(\theta').$$

其次，我们来证明(ii)⇒(i)。

由条件有

$$(\mathscr{L}\mathscr{L})(\theta)=\bigvee_{uv=\theta}[\mathscr{L}(u)\wedge\mathscr{L}(v)]$$

$$\leqslant\bigvee_{uv=\theta}\mathscr{L}(uv)=\mathscr{L}(\theta)$$

从而$\mathscr{L}\mathscr{L}\leqslant\mathscr{L}$成立，又$\mathscr{L}_\Lambda\leqslant\mathscr{L}$，从而$\widehat{\mathscr{L}}\leqslant\mathscr{L}$，反之$\mathscr{L}\leqslant\widehat{\mathscr{L}}$是显然的，故有$\widehat{\mathscr{L}}=\mathscr{L}$.

我们从上所述可以给出 X 内的不分明语言与不分明代数结构之间一个简单结果。

推论 1 不分明语言$\mathscr{L}$是闭的($\widehat{\mathscr{F}}=\mathscr{L}$)，当且仅当$\mathscr{L}$是 X^*的不分明子单半群(Submonoid).

证 我们在研究不分明集的代数结构时，曾经定义：若不分明集合 $\underset{\sim}{A}\in\mathscr{F}_L(X)$ 满足如下的条件，就叫做在运算"·"下是封闭的。

$$\underset{\sim}{A}(x\cdot y)\geqslant\underset{\sim}{A}(x)\wedge\underset{\sim}{A}(y),\quad \forall x,\forall y\in X.$$

如果$(X,\cdot)$是群，则 X 的不分明子群是满足如下性质的 $\underset{\sim}{A}\in\mathscr{L}_L(X)$。

$$\underset{\sim}{A}(x,y)\geqslant\underset{\sim}{A}(x)\wedge\underset{\sim}{A}(y),\qquad \forall x,\ \forall y\in X;$$
$$\underset{\sim}{A}(x^{-1})\geqslant\underset{\sim}{A}(x)\qquad \forall x\in X.$$

由该定义可知，定理中的第二个不等式即表示$\mathscr{L}$为 X^*的不分明子单半群，故推论成立。

5.1.4 语言变量

在人类的语言中有一类词直接涉及数值. 如[大]、[小]、[长]、[短]、[轻]、[重]、[多少]、[年令]等，它们的论域直接

为 $R=(-\infty,+\infty)$或 R 的子集。对于这种特殊的词，我们引入语言变量这一概念。

语言变量，简单地说，是一个变量，其取值为人类语言中的词或句子。

例如，“年令”这个词，若解释为语言变量，则年令的术语集是

$$
\begin{aligned}
T(\text{年令}) = &\text{年轻}+\text{年老}+\text{很年轻}+\text{不年轻}\\
&+\text{很老}+\text{非常非常年轻}\\
&+\text{比较年轻}+\text{有几分年轻}+\cdots
\end{aligned}
$$

其中每一项是人类语言(如 $U=[0,100]$)的不分明子集的一个标签。

又如，若在 $X=\{1,2,3,4,5\}$上定义了

$$[\text{大}]=0.5/4+1/5,$$

$$[\text{小}]=1/1+0.2/2.$$

则

$$[\text{很大}]=H_2[\text{大}]=0.25/4+1/5,$$

$$[\text{很小}]=H_2[\text{小}]=1/1+0.04/2.$$

它们亦是语言变量。

显见，语言变量是一个集函数。它连结两个值：(a)一个造句法的规律，它在 $T(x)$中定义为合式句(Well-formed Sentence)，(b)关于词意的值，用 $T(x)$中的项来确定。若 x 是 $T(x)$中一项，则其意是 U 的一个子集。$T(x)$中一个原始项，表示原始不分明集，即是一个早先定义过的项，并且作为 $T(x)$中非原始项计算的基础。例如，在我们上面所举的例子中，原始项是“年轻”和“年老”，它的含义可以由相应的隶属函数来定义。则其它非原始项的含义，均由原始项的隶

属函数经过词义结构来计算。如

$$[很年轻]=H_2[年轻]=[年轻]^2$$

即

$$\mu_{[很年轻]}=(\mu_{[年轻]})^2 .$$

又

$$[不很年轻]=\overline{[很年轻]}$$

即

$$\mu_{[不很年轻]}=1-\mu_{[很年轻]}$$

$$=1-(\mu_{[年轻]})^2.$$

语言变量是不分明集，它的值域是数空间。对于同一个论域上的语言变量，可以定义它们之间的运算。

设 α,β 为两个语言变量，定义

$$(\alpha*\beta)(z)\triangleq\bigvee_{x*y=z}(\alpha(x)\wedge\beta(y)).$$

例 5.1.3 设

$$\alpha=1/1+0.8/2,$$
$$\beta=0.2/2+0.8/3,$$

则

$$\begin{aligned}\alpha+\beta&=1\wedge0.2/(1+2)+1\wedge0.8/(1+3)\\&\quad+0.8\wedge0.2/(2+2)+0.8\wedge0.8/(2+3)\\&=0.2/3+0.8/4+0.2/4+0.8/5\\&=0.2/3+0.8\vee0.2/4+0.8/5\\&=0.2/3+0.8/4+0.8/5;\end{aligned}$$

$$\begin{aligned}\alpha-\beta&=1\wedge0.2/(1-2)+1\wedge0.8/(1-3)\\&\quad+0.8\wedge0.2/(2-2)+0.8\wedge0.8/(2-3)\\&=0.2/(-1)+0.8/(-2)+0.2/0+0.8/(-1)\\&=0.2/0+0.2\vee0.8/(-1)+0.8/(-2)\end{aligned}$$

$$= 0.2/0 + 0.8/(-1) + 0.8/(-2);$$

$$\alpha \cdot \beta = 1 \wedge 0.2/1 \cdot 2 + 1 \wedge 0.8/1 \cdot 3 + 0.2 \wedge 0.8/2 \cdot 2 + 0.8 \wedge 0.8/2 \cdot 3$$

$$= 0.2/2 + 0.8/3 + 0.2/4 + 0.8/6;$$

$$\alpha \div \beta = 1 \wedge 0.2/1 \div 2 + 1 \wedge 0.8/1 \div 3 + 0.2 \wedge 0.8/2 \div 2 + 0.8 \wedge 0.8/2 \div 3$$

$$= 0.2\Big/\frac{1}{2} + 0.8\Big/\frac{1}{3} + 0.2/1 + 0.8\Big/\frac{2}{3}$$

由此可以看出，对语言变量的四则运算和不分明数的四则运算是一样的。可见语言变量是一个 F 数(不分明数)。

§2 语言概率

本节我们来讨论语言学中的概率问题。

在通常的概率论中，概率是取值于[0,1]的一个集函数。其涉及的事件是样本空间 Ω 的某些子集所构成的 σ 代数中的元素，这是 Ω 的一个精确子集。在讨论不分明概率时，涉及的是 Ω 的某些不分明子集所构成的不分明事件域中之元素。这里考虑的是不分明集合。但两者所讨论的概率却不存在模糊性。

值得注意的是，在语言学中需要考查诸如"概率很大"，"概率很小"，"概率不大"等不分明概念。由前所述，它应当是一个语言意义下的概念，即是语言意义下的概率，论域为[0,1]或[0,1]的一个子集的语言变量。而且在论域中的位置未加以明确地限定，可以是不分明的。

定义5.2.1 设$\mathscr{F}$([0,1])表示以[0.1]为论域的全体不分明子集所构成的集族。它的某一选定的子集族 ε，称为语

言概率的值空间。ε 中的元素称为概率语言变量。

语言概率记作$\underset{\sim}{P}$。

过去的概率P,是 Ω 的 Borel 域$\mathscr{B}$到[0,1]的一个映射．而语言概率 $\underset{\sim}{P}$，则是从$\mathscr{B}$到 ε 的一个映射。

为了研究$\underset{\sim}{P}$，首先必需研究清楚 ε。我们希望 ε 应 满足两方面的要求。第一，它应具有一定的语言特点．应该是一个不分明语言系统，包括若干单词，而且应该关于逻辑运算和算子作用封闭。其次，它能适应概率运算的要求。

5.2.1 语言概率值空间的语义结构

首先我们认为 ε 应当包含以下几种单词：

(1) “p”(p为 [0,1]中一实数)

这就是说，不分明概率的值域空间应当包含普通概率的值域空间。

“p”的隶属函数规定为

$$p(q)\triangleq\delta(p,q)=\begin{cases}1, & \text{当 } p=q,\\ 0, & \text{当 } p\neq q.\end{cases}\quad (q\in[0,1])\tag{5.2.1}$$

(2) “很可能”

我们规定“很可能”的隶属函数为

$$\text{“很可能”}(p)=\begin{cases}0, & 0\leqslant p\leqslant a,\\ 2\left(\dfrac{p-a}{1-a}\right)^2, & 0\leqslant p\leqslant\dfrac{a+1}{2},\\ 1-2\left(\dfrac{p-a}{1-a}\right)^2, & \dfrac{a+1}{2}\leqslant p\leqslant 1.\end{cases}\tag{5.2.2}$$

此处 a 为一参数，$a>\dfrac{1}{2}$。

在实际问题中，往往遇到的是离散情形。从而也可以将

它的支集缩小到仅仅包含有限个点，也可说是把论域[0,1]换为V。通常V取为

$$V=\{0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1\}$$

而此时，常规定

$$\text{“很可能”}=0.5/0.6+0.7/0.7+0.9/0.8+1/0.9+1/1. \tag{5.2.3}$$

(3)“很不可能”

我们规定“很不可能”(p)=“很可能”$(1-p)$。

如果我们取“很可能”的隶属函数为上面所述的规定(2),则

$$\text{“很不可能”}(p)=\begin{cases}0, & 0\leqslant 1-p\leqslant a,\\ 2\left(\dfrac{1-p-a}{1-a}\right)^2, & a\leqslant 1-p\leqslant \dfrac{a+1}{2},\\ 1-2\left(\dfrac{1-p-a}{1-a}\right)^2, & \dfrac{a+1}{2}\leqslant 1-p\leqslant 1.\end{cases}$$

若取为(3)的规定，则

$$\begin{aligned}\text{“很不可能”}(p)&=0.5/(1-0.6)+0.7/(1-0.7)\\&\quad+0.9/(1-0.8)+1/(1-0.9)+1/(1-1)\\&=1/0+1/0.1+0.9/0.2+0.7/0.3+0.5/0.4.\end{aligned}$$

由以上这些词(原始词)，经过逻辑运算，可得到一系列新的语言概率词。例如

$$\begin{aligned}\text{“不很可能”}(p)&=[(\text{很可能})^c](p)\\&=1-\text{“很可能”}(p).\end{aligned}$$

$$\begin{aligned}&\text{“很可能或很不可能”}(p)\\&\quad=\text{“很可能”}(p)\vee\text{“很不可能”}(p).\end{aligned}$$

另外，由原始词经过算子运算，亦得到一些其它词。

例如：

$$“稍许可能”(p)=[稍许][可能](p)$$

$$=H_{\frac{1}{2}}[可能](p)=[可能(p)]^{1/2}$$

$$=\begin{cases}0, & 0\leqslant p\leqslant a,\\ \sqrt{2}\left(\frac{p-a}{1-a}\right), & a\leqslant p\leqslant\frac{1+a}{2},\\ \sqrt{1-2\left(\frac{p-a}{1-a}\right)^2}, & \frac{a+1}{2}\leqslant p\leqslant 1.\end{cases}$$

$$“非常可能”(p)=([得可能](p))^2$$

$$=\begin{cases}0, & 0\leqslant p\leqslant a,\\ 4\left(\frac{p-a}{1-a}\right)^4, & a\leqslant p\leqslant\frac{a+1}{2},\\ \left[1-2\left(\frac{p-a}{1-a}\right)^2\right]^2 & \frac{a+1}{2}\leqslant p\leqslant 1.\end{cases}$$

$$“差不多是\ p”(q)=F[p(q)]$$

$$=\begin{cases}e^{-(p-q)^2}, & |q-p|\leqslant\delta,\\ 0, & |q-p|>\delta.\end{cases}\quad (q\in[0,1])$$

“差不多是 0”=“几乎不可能”(p)；

“差不多是 1”=“几乎一定(发生)”.

$$“偏向于很可能”(p)=d_{\frac{1}{2}}[很可能](p)$$

$$=\begin{cases}0, & p<\frac{1+a}{2},\\ 1, & p>\frac{1+a}{2},\end{cases}$$

$$“偏向于很不可能”(p)=d_{\frac{1}{2}}[很不可能](p)$$

$$= \begin{cases} 1, & p < \dfrac{1-a}{2}, \\ 0, & p > \dfrac{1-a}{2}. \end{cases}$$

通过以上所述，可以看出语言概率值空间的语义结构.

5.2.2 ε 中的运算关系

我们希望概率语言变量经过四则运算之后，仍应是概率语言变量。可是实际情形并不如此，若按前面规定的语言运算规则计算，所得的结果可以不是概率语言变量。于是 Zadeh 提出了概率语言变量的运算法则。

定义 5.2.2 设 $\pi_1, \pi_2, \cdots, \pi_n \in \varepsilon$。$a_1, a_2, \cdots, a_n \in [0,1]$，则 $\{\pi_i\}$ 的线性组合被定义为一个不分明语言变量：

$$(a_1\pi_1 + a_2\pi_2 + \cdots + a_n\pi_n)(p)$$

$$\triangleq \left[\bigvee_{(a_1\pi_1 + \cdots + a_n\pi_n = p)} (\pi_1(p_1) \wedge \cdots \wedge \pi_n(p_n) \right] \div \mu$$

此处

$$\mu = \bigvee_{p_1 + p_2 + \cdots + p_n = 1} [\pi_1(p_1) \wedge \cdots \wedge \pi_n(p_n)].$$

约定空集的上确界为 0，则当 $p>1$ 时，隶属函数为 0。故 $\pi_1, \pi_2, \cdots, \pi_n$ 的线性组合仍将支集限制在 $[0,1]$ 之内。这就是说，概率语言变量对于如此规定的运算是封闭的。

例 5.2.1 考虑样本空间仅有两个元素的情形。设

$\pi_1 =$ “很可能”，

$\pi_2 =$ “很不可能”。

由 π_1，π_2 的规定，有

$$\pi_1(p) = \pi_2(1-p). \tag{5.2.4}$$

所以

$$(a_1\pi_1+a_2\pi_2)(p)$$

$$=\bigvee_{\substack{a_1p_1+a_2p_2=p\\ p_1+p_2=1}}(\pi_1(p_1)\wedge\pi_2(p_2))]+\mu$$

$$=\frac{1}{\mu}\bigvee_{a_1p_1+a_2(1-p_1)=p}(\pi_1(p_1)\wedge\pi_2(1-p_1))$$

$$=\frac{1}{\mu}\bigvee_{a_1p_1+a_2(1-p_1)=p}(\pi_1(p_1))$$

这是由于 $\pi_2(1-p_1)=\pi_1(p_1)$ 之故。再根据 p_1 的限制有

$$p_1=\frac{p-a_2}{a_1-a_2},$$

故

$$(a_1\pi_1+a_2\pi_2)(p)$$

$$=\frac{1}{\mu}\pi_1\left(\frac{p-a_2}{a_1-a_2}\right)$$

$$=\begin{cases}\frac{1}{\mu}\pi_1\left(\frac{p-a_2}{a_1-a_2}\right), & p\in(a_1,a_2)\\ 0, & \text{其它}\end{cases}$$

但是

$$\mu=\bigvee_{p_1+p_2=1}(\pi_1(p_1)\wedge\pi_2(p_2))=\bigvee_{p_1\in V}\pi_1(p_1)=1.$$

故

$$(a_1\pi_1+a_2\pi_2)(p)$$

$$=\begin{cases}\pi_1\left(\frac{p-a_2}{a_1-a_2}\right) & p\in(a_1,a_2)\\ 0, & \text{其它}\end{cases}\qquad(5.2.5)$$

按上节之规定设“很可能”的隶属函数为

$$\text{“很可能”}(p)=\begin{cases}0, & 0\leqslant p\leqslant a,\\ 2\left(\dfrac{p-a}{1-a}\right)^2, & a\leqslant p\leqslant\dfrac{1+a}{2},\\ 1-2\left(\dfrac{p-a}{1-a}\right)^2, & \dfrac{1+a}{2}\leqslant p\leqslant 1.\end{cases}$$

不妨取 $a=0.8$，将其代入上式，由（5.2.5）有：当 $0.3\leqslant p\leqslant 0.7$ 时，

$$(a_1\pi_1+a_2\pi_2)(p)=\pi_1\left(\frac{p-0.3}{0.4}\right)$$

$$=\pi_1\left(\frac{p-0.3}{0.7-0.3}\right)$$

$$=\begin{cases}0 & 0\leqslant\dfrac{p-0.3}{0.4}<0.8,\\ 2\left[\left(\dfrac{p-0.3}{0.4}-0.8\right)\quad(1-0.8)\right]^2, & \\ & 0.8\leqslant\dfrac{p-0.3}{0.4}<0.9,\\ 1-2\left[\left(\dfrac{p-0.3}{0.4}-0.8\right)\Big/(1-0.8)\right)^2, & \\ & 0.9\leqslant\dfrac{p-0.3}{0.4}\leqslant 1.\end{cases}$$

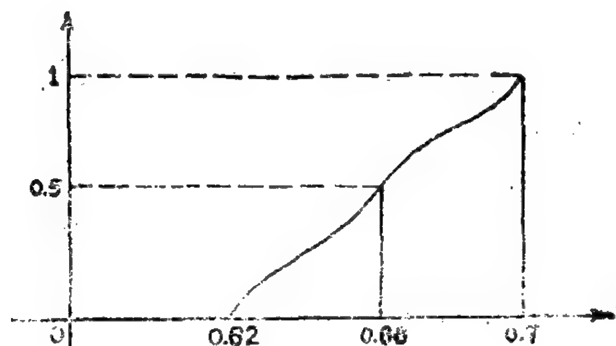
$$=\begin{cases}0, & 0.3\leqslant p<0.62,\\ 2\left(\dfrac{p-0.32}{0.08}\right)^2, & 0.62\leqslant p<0.66,\\ 1-2\left(\dfrac{p-0.62}{0.08}\right)^2, & 0.66\leqslant p\leqslant 0.7.\end{cases}$$

当 $p\in[0.3,\ 0.7]$时，$(a_1\pi_1+a_2\pi_2)(p)=0$。故得

$$[0.7(\text{“很可能”})+0.3(\text{“很不可能”})](p)$$

$$
= \begin{cases} 0 & 0 \leqslant p < 0.62, \\ 2\left(\dfrac{p-0.62}{0.08}\right)^2 & 0.62 \leqslant p < 0.66, \\ 1-2\left(\dfrac{p-0.62}{0.08}\right)^2, & 0.66 \leqslant p < 0.7, \\ 0, & 0.7 \leqslant p \leqslant 1. \end{cases}
$$

其图形如下：



0.7“很可能”+0.3“很不可能”的隶属函数图形

图 5-2-1

注 当我们有了(5.2.5)式之后，求隶属度，无异于在线性约束条件下求解一个非线性规划问题。其线性约束条件是

$$a_1 p_1 + a_2 p_2 + \cdots + a_n p_n = p,$$

$$p_1 + p_2 + \cdots + p_n = p.$$

而极大化函数为

$$\pi_1(p_1) \wedge \pi_2(p_2) \wedge \cdots \wedge \pi_n(p_n).$$

5.2.3 语言概率的数学模型

设(Ω, F, P)为一概率空间，又设$A \in F$。而且假定Ω为有限的：

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}.$$

又给定了概率语言变量系 ε，从中指定一个序列

$$\pi_1, \pi_2, \cdots, \pi_n.$$

定义

$$\underset{\sim}{P}(\omega_i) \triangleq \pi_i \quad (i=1,2,\cdots,n)$$

且令

$$\underset{\sim}{P}(A) = a_1\pi_1 + a_2\pi_2 + \cdots + a_n\pi_n,$$

其中

$$a_i = \begin{cases} 1, & \omega_i \in A, \\ 0, & \omega_i \overline{\in} A. \end{cases} \quad (i=1,2,\cdots,n)$$

易见映射 $\underset{\sim}{P}$: $F \to \varepsilon$ 具有以下性质：

(1) 有限可加性，即

对任意 $A_1, A_2 \in F$，$A_1 \cap A_2 = \phi$，则

$$\underset{\sim}{P}(A_1 \cup A_2) = \underset{\sim}{P}(A_1) + \underset{\sim}{P}(A_2).$$

此处 $\underset{\sim}{P}(A_1) + \underset{\sim}{P}(A_2)$ 的意思是指：若

$$\underset{\sim}{P}(A_1) = a_{11}\pi_1 + a_{12}\pi_2 + \cdots + a_{1n}\pi_n,$$

$$\underset{\sim}{P}(A_2) = a_{21}\pi_1 + a_{22}\pi_2 + \cdots + a_{2n}\pi_n,$$

则

$$\underset{\sim}{P}(A_1) + \underset{\sim}{P}(A_2) = (a_{11} + a_{21})\pi_1 + (a_{12} + a_{22})\pi_2 + \cdots + (a_{1n} + a_{2n})\pi_n.$$

(2) 正规性，即

$$\underset{\sim}{P}(\Omega) = 1, \quad (1 \text{ 是}\{1\}\text{的隶属函数}).$$

由这些性质可见，$\underset{\sim}{P}$ 是一个概率，为此

定义 5.2.2 如上建立的 $(\Omega, F, \varepsilon, \underset{\sim}{P})$ 叫做一个离散的语言概率场。$\underset{\sim}{P}$ 叫做语言概率。

注 关于语言概率问题在这里需要再系统的解释一下。

设(Ω,F,P)是一个分明的概率空间。若$A\in F$，则$P(A)$应取值于$[0,1]$上的一个值。要是$P(A)$未知，那么它就是取值于$[0,1]$上的一个数值变量。将事件A与语言变量相联系，用$\underset{\sim}{P}(A)$来表示，并以$P(A)$作为它的基础变量。即分配到$\underset{\sim}{P}(A)$的每一个值为$P(A)$上的一个不分明关系。如此，与每个事件A一起，我们连带一个在$\widetilde{\mathscr{P}}([0,1])$的子集中取值的参数，包含在一个(可列)集$\varepsilon\subseteq\widetilde{\mathscr{P}}([0,1])$中取值的$\underset{\sim}{P}(A)$中，其中$\varepsilon$的每一个元素是$[0.1]$的一个不分明子集，而它属于一个词集$T(p)$。

$\underset{\sim}{P}(A)=I$，$I\in\varepsilon$，表示I是$P(A)$上的一个不分明关系，并且称I为语言概率值。事实上，$\underset{\sim}{P}$是一个均值，被认为是F到ε的一个多值映射。

5.2.4 F 事件的语言概率

无论从逻辑系统，还是从语言本身，都需要我们考虑不分明事件的语言概率。

人类为了相互交流思想、描述客观事物，在客观实践中已逐步形成了一套结构细密的语言系统。其中不仅有不分明语言，涉及语言概率的语言，而且还有一些语言表述涉及不分明事件的语言概率。例如，“明天是个好天气的可能性很大”；“你要找的人很可能是个高个子”；“这个方案受到绝大多数人拥护的可能性不大”等等。虽然它们每一个叙述得很不确切，但是谈话的人能相互理解而且比较直观。

我们先来分析一下上面所举的语句。在上面语句中涉及到“好天气”、“高个子”、“大多数”等概念。由以往知识知道，这些是些不分明事件。而涉及概率的却是“很大”、“很”、“不大”等，又是不分明的。因而这里就必须考虑不分明事件的

不分明概率问题，亦即不分明事件的语言概率问题。

设 Ω 为样本空间，$\Omega=\{\omega_i\}$，$i=1,2,\cdots,n$。$\underset{\sim}{A}$ 为 Ω 上的不分明子集，由隶属函数 $\mu_{\underset{\sim}{A}}(\omega_i)$所确定。$\underset{\sim}{F}$ 为 Ω 生成的不分明事件域。

定义 5.2.3 在语言概率场$(\Omega,\underset{\sim}{F},\varepsilon,\underset{\sim}{P})$上，$\underset{\sim}{A}\in\underset{\sim}{F}$，则定义

$$\underset{\sim}{P}(\underset{\sim}{A})=\mu_{\underset{\sim}{A}}(\omega_1)\pi_1+\mu_{\underset{\sim}{A}}(\omega_2)\pi_2+\cdots+\mu_{\underset{\sim}{A}}(\omega_n)\pi_n$$

为 $\underset{\sim}{A}$ 的语言概率，记作 $\underset{\sim}{P}(\underset{\sim}{A})$。其中 $\pi_i=\underset{\sim}{P}(\omega_i)$，$i=1,2,\cdots,n$。

由定义可见，$\underset{\sim}{P}(\underset{\sim}{A})$实质上是加权平均。

例 5.2.2 设

$$\Omega=\{a,b,c\}$$

$$\underset{\sim}{A}=0.4/a+1/b+0.8/c;$$

$$\underset{\sim}{P}(a)=\pi_a=\text{“接近于 0.3”}=0.6/0.2+1/0.3+0.6/0.4;$$

$$\underset{\sim}{P}(b)=\pi_b=\text{“接近于 0.6”}=0.6/0.5+1/0.6+0.6/0.7;$$

$$\underset{\sim}{P}(c)=\pi_c=\text{“接近于 0.1”}=0.6/0+1/0.1+0.6/0.2.$$

则不分明事件 $\underset{\sim}{A}$ 的语言概率为

$$\begin{aligned}\underset{\sim}{P}(\underset{\sim}{A})=&0.4(0.6/0.2+1/0.3+0.6/0.4)\\&+1(0.6/0.5+1/0.6+0.6/0.7)\\&+0.8(0.6/0+1/0.1+0.6/0.2)\end{aligned}$$

此即$\underset{\sim}{P}(\underset{\sim}{A})=0.4\pi_a+1\pi_b+0.8\pi_c$，为 π_a,π_b,π_c 的线性组合。故按 ε 中运算之定义，应为适合条件

$$0.4p_1+p_2+0.8p_3=p;$$

$$p_1 + p_2 + p_3 = 1$$

的项。

上式括号中的"+"为 Zadeh 记号，而括号外的 记 号 为 +运算。按照运算定义，即将以下三括号逐项搭配：先将第一个括号 $0.4\pi_a$ 与 π_b 搭配，得项：

$0.6/0.4\times0.2+0.5$； $0.6/0.4\times0.2+0.6$；

$0.6/0.4\times0.2+0.7$； $0.6/0.4\times0.3+0.5$；

$1/0.6\times0.3+0.6$； $0.6/0.4\times0.3+0.7$；

$0.6/0.4\times0.4+0.5$； $0.6/0.4\times0.4+0.6$；

$0.6/0.4\times0.4+0.7$。

再将其与第三括号

$$0.6/0+1/0.8\times0.1+0.6/0.8\times0.2$$

搭配，共得 27 项：

$0.6/0.4\times0.2+0.5$； $0.6/0.4\times0.2+0.6$；

$0.6/0.4\times0.4+0.7$； $0.6/0.4\times0.2+0.7$；

$0.6/0.4\times0.3+0.5$； $1/0.4\times0.3+0.6$；

$0.6/0.4\times0.3+0.7$； $0.6/0.4\times0.4+0.5$；

$0.6/0.4\times0.4+0.6$；

$0.6/0.4\times0.2+0.5+0.8\times0.1$；

$0.6/0.4\times0.2+0.5+0.8\times0.2$；

$0.6/0.4\times0.2+0.6+0.8\times0.1$；

$0.6/0.4\times0.2+0.6+0.8\times0.2$；

$0.6/0.4\times0.2+0.7+0.8\times0.1$；

$0.6/0.4\times0.2+0.7+0.8\times0.2$；

$0.6/0.4\times0.3+0.5+0.8\times0.1$；

$0.6/0.4\times0.3+0.5+0.8\times0.2$；

$1/0.4\times0.3+0.6+0.8\times0.1;$

$1/0.4\times0.3+0.6+0.8\times0.2;$

$0.6/0.4\times0.3+0.7+0.8\times0.1;$

$0.6/0.4\times0.3+0.7+0.8\times0.1;$

$0.6/0.4\times0.4+0.5+0.8\times0.1;$

$0.6/0.4\times0.4+0.5+0.8\times0.2;$

$0.6/0.4\times0.4+0.6+0.8\times0.1;$

$0.6/0.4\times0.4+0.6+0.8\times0.2;$

$0.6/0.4\times0.4+0.7+0.8\times0.1;$

$0.6/0.4\times0.4+0.7+0.8\times0.1.$

其中满足 $p_1+p_2+p_3=1$ 者有以下诸项：

$1/0.4\times0.3+0.6+0.8\times0.1$

$0.6/0.4\times0.3+0.7;$

$0.6/0.4\times0.4+0.6;$

$0.6/0.4\times0.2+0.6+0.8\times0.2;$

$0.6/0.4\times0.2+0.7+0.8\times0.1;$

$0.6/0.4\times0.3+0.5+0.8\times0.2;$

$0.6/0.4\times0.4+0.5+0.8\times0.1;$

因此，将以上各项按 Zadeh 记号"+"起来，得

$$
\begin{aligned}
P(A) = & 0.6/(0.4\times0.3+0.7)+0.6/(0.4\times0.4+0.6 \\
& +0.6/(0.4\times0.2+0.6+0.8\times0.2) \\
& +0.6/(0.4\times0.2+0.7+0.8\times0.1) \\
& +0.6/(0.4\times0.3+0.5+0.8\times0.2) \\
& +0.6/(0.4\times0.4+0.5+0.8\times0.1) \\
& +1/(0.4\times0.3+0.6+0.8\times0.1) \\
= & 0.6/0.82+0.6/0.76+0.6/0.84+0.6/0.86
\end{aligned}
$$

$+0.6/0.78+0.6/0.74+1/0.80$

若采用 Zadeh 记号：

$$\frac{\text{隶属度}\ \mu}{\text{甲}\oplus\text{乙}}=\frac{\text{隶属度}\ \mu}{\text{甲}}+\frac{\text{隶属度}\ \mu}{\text{乙}}$$

进行化简，即将隶属度一样的归并在一起。于是有

$$\underset{\sim}{P}(\underset{\sim}{A})=0.6/(0.74\oplus0.76\oplus0.78\oplus0.82\oplus0.84\oplus0.86)+1/0.80.$$

即不分明事件 $\underset{\sim}{A}$ 的 F 概率接近于 0.8。

注 最后我们提出 Nguytn 在 1977 年引进的语言随机变量及语言期望。

设 ξ 是一个随机变量，取值于一有限集合 $U=\{u_1,u_2,\cdots,u_n\}\subseteq\boldsymbol{R}$ 中。又设 $\xi=u_i$ 的概率为

$$P(\xi=u_i)=p_i,\quad i=1,2,\cdots,n.$$

考虑语言变量 $\underset{\sim}{P}_i$，显然它是一个语言概率。连带 ξ 每一个 n 重 $(A_1,A_2,\cdots,A_n)$ 构成一个语言概率分布表.其中 A_i 是分配于 $\underset{\sim}{P}_i$ 上的语言值。这样的表，我们称为语言概率分布。

随机变量 ξ，连带它的语言概率分布，称为语言随机变量.

值得提出，对于 $(A_1,A_2,\cdots,A_n)$，可有构造上的差别。

例 5.2.3 ξ 的语言均匀分布可以表示为

$$A_i=\text{“}\frac{1}{n}\text{”}\quad i=1,2,\cdots,n$$

其中“$\frac{1}{n}$”是 $[0,1]$ 上的不分明子集，标明“接近于 $\frac{1}{n}$”。

设 ξ 是一个在 $U=\{u_1,u_2,\cdots,u_n\}\subseteq\boldsymbol{R}$ 上取值的语言随机变量，具有语言概率分布 $(A_1,A_2,\cdots,A_n)$。

定义 5.2.4 称

$$u_1A_1+u_2A_2+\cdots+u_nA_n$$

为 ξ（关于 $(A_1,A_2,\cdots,A_n)$）的语言期望。记作 $\underset{\sim}{E}(\xi)$。

$\underset{\sim}{E}(\xi)$的含义，可以从开拓原理演绎出来。特别注意，[0,1]上的每一个 F 子集，首先是变量 p_i 上的一个 F 限制(restriction)。因此，A_i 之间的相互影响是 p_i 间存在着约束，表示为

$$S\subseteq[0,1]^n,\quad S=\{(p_1,\cdots,p_n)\mid p_1+p_2+\cdots+p_n=1\}.$$

若 A_i 是非相互影响的，则被 $(A_1,A_2,\cdots,A_n)$ 包含的限制由

$$\mu_{(A_1,A_2,\cdots,A_n)}(x_1,x_2,\cdots,x_n)=\left[\bigwedge_{i=1}^{n}\mu_{A_i}(x_i)\right]\cap \lg(x_1,\cdots,x_n)$$

给出。由开拓原理可见 $\underset{\sim}{E}(x)$ 是实直线上的 F 集合，由下式给定

$$\mu_{\underset{\sim}{E}(X)}(z)=\bigvee_{\substack{(x_1,\cdots,x_n)\in[0,1]^n\\ x_1+\cdots+x_n=1\\ u_1x_1+\cdots+u_nx_n=z}}\left[\bigvee_{i=1}^{n}\mu_{A_i}(x_i)\right]$$

关于 $\underset{\sim}{E}(X)$ 的计算，即确定它的隶属函数问题，转化成为一个具有线性限制的一个非线性规划的求解问题。

§3　F 事件与 F 概率

F 事件的语言概率是 F 事件的 F 概率。这一节我们介绍 Kwakernaak 的观点.他的文章和 Nguytn 的发表在同一年，基本论点相似。

设 $\boldsymbol{X}$ 是一个 KF 随机变量，A 是 R 中一个 Borel 集。则

F概率$P_r(\boldsymbol{X}\in A)$定义为 $\boldsymbol{X}=(\widetilde{\mathscr{X}},X)$ 在映射$\widetilde{U}\to\widetilde{\mathscr{P}}$下$R$ 中的像，其中$\widetilde{U}\in A$.所以我们有$P_r(\boldsymbol{X}\in A)=(R,(P_r(\boldsymbol{X}\in A))$，而

$$(P_r(\boldsymbol{X}\in A))=\sup_{\widetilde{U}\in\widetilde{\mathscr{X}}:\widetilde{P}(\widetilde{U}\in A)=p}\ \inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}}(X_\omega(\widetilde{U}(\omega,\omega'))).$$

因为 $0\leqslant\widetilde{P}(\widetilde{U}\in A)\leqslant 1$，对于 $p\notin[0,1]$有$(P_r(\boldsymbol{X}\in A))=0$。这有利于把 $P_r(\boldsymbol{X}\in A)$考虑为 $\boldsymbol{X}$ 在两个映射下的像。记

$$\widetilde{P}(\widetilde{U}\in A)=\int d\mathscr{P}(\omega)\int_{\omega':\widetilde{U}(\omega,\omega')\in A}d\mathscr{P}'(\omega'),$$

而第一个映射：$\widetilde{\mathscr{X}}\to\mathscr{X}$，由

$$\widetilde{U}\to\int_{\omega':\widetilde{U}(\cdot,\omega')\in A}d\mathscr{P}'(\omega')$$

确定。第二个映射：$\mathscr{X}\to R$ 由 $U\to EU$ 确定。

定理5.3.1 $\boldsymbol{X}=(\widetilde{\mathscr{X}},X)$在 $\mathscr{X}$ 中的像，由

$$\widetilde{U}\to\int_{\omega':\widetilde{U}(\cdot,\omega')\in A}dP'(\omega') \tag{3.1}$$

给定。即 $I^{\boldsymbol{X}\in A}=(\mathscr{X},I^{\boldsymbol{X}\in A})$，而

$$I_\omega^{\boldsymbol{X}\in A}(\boldsymbol{\pi})=\sup_{U'\in\mathscr{X}':\mathscr{P}'(U'\in A)=\pi}\ \inf_{\omega'\in\Omega'}X_\omega(U'(\omega)),\quad\boldsymbol{\pi}\in\bar{R}.$$

其中 $\mathscr{X}'$ 是定义在$(\Omega',\mathscr{F}',\mathscr{P}')$上一切随机变量的集合。

证 令 $V\in\mathscr{X}$，则在映射(3.1)作用下 $\boldsymbol{X}$ 的像中 V 的隶属度

$$\sup_{\widetilde{U}\in\widetilde{\mathscr{X}}:(\forall\omega)\mathscr{P}'(\widetilde{U}(\omega,\cdot)\in A)=V(\omega)} \inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}} X_\omega(\widetilde{U}(\omega,\omega'))$$

$$=\inf_{\omega\in\Omega}\sup_{U\in\mathscr{X}:\mathscr{P}'(U'\in A)=V(\omega)}\inf_{\omega'\in\Omega'} X_\omega(U'(\omega'))$$

从而得证。

定理 5.3.2 记

$$r'_A(\omega)=\sup_{x\in A}X_\omega(x),\quad r''_A(\omega)=\sup_{x\in A^c}X_\omega(x).$$

则 $I^{\boldsymbol{X}\in A}$(称为 $\boldsymbol{X}\in A$ 的标指函数)由下式确定

$$I^{\boldsymbol{X}\in A}_\omega(\pi)=\begin{cases} r'_A(\omega) & 若\ \pi=0\\ \min[r'_A(\omega),r''_A(\omega)], & 若\ 0<\pi<1.\\ r''_A(\omega). & 若\ \pi=1.\end{cases}$$

并且为单峰的。

证 若记 $I^{\boldsymbol{X}\in A}_\omega(\pi)=\sup_{U'\in\mathscr{X}':\mathscr{P}(U'\in A)=\pi}$

$$\min\Big[\inf_{\omega':U'(\omega')\in A}X_\omega(U'(\omega')),$$

$$\inf_{\omega':U'(\omega')\in A^c}X_\omega(U'(\omega'))\Big]$$

分别考虑 $\pi=0$, $0<\pi<1$ 及 $\pi=1$ 的情形，则结论显然。

若 $\boldsymbol{X}=(\widetilde{\mathscr{X}},X)$是一个非 F 随机变量，则$I^{\boldsymbol{X}\in A}$即通常的示性函数。

定理 5.3.3 $P_r(\boldsymbol{X}\in A)=EI^{\boldsymbol{X}\in A}$

证 ∵$\widetilde{P}(\widetilde{U}\in A)$是由(3.1)式确定的映射$\widetilde{\mathscr{X}}\to\mathscr{X}$ 和由 $U\to EU$ 确定的映射 $\mathscr{X}\to R$ 的复合。由定理 5.3.1 即得证。

设 $\boldsymbol{X}=(\widetilde{\mathscr{X}},X)$和 $\boldsymbol{Y}=(\widetilde{\mathscr{Y}},Y)$是两个 KF 随机变量，A 和

B 是 R 上的两个 Borel 集。那么定义联合概率 $P_r(\boldsymbol{X}\in A, \boldsymbol{Y}\in B)$ 为一 F 数，它有隶属函数

$$(P_r(\boldsymbol{X}\in A,\boldsymbol{Y}\in B))(p) = \sup_{\substack{\widetilde{U}\in\widetilde{\mathscr{X}},\widetilde{V}\in\widetilde{\mathscr{Y}}:\\ \widetilde{P}(\widetilde{U}\in A,\widetilde{V}\in B)=p}} \inf_{\substack{\omega\in\Omega\\ \omega'\in\Omega'}} \min[X_\omega(\widetilde{U}(\omega,\omega')), Y_\omega(V(\omega,\omega'))],\ p\in[0,1].$$

类似定理 5.3.3 我们有

$$P_r(\boldsymbol{X}\in A,\boldsymbol{Y}\in B)=EI^{\boldsymbol{X}\in A,\boldsymbol{Y}\in B}.$$

其中 $I^{\boldsymbol{X}\in A,\boldsymbol{Y}\in B}(\pi)$

$$= \sup_{U',V'\in\mathscr{X}':P'(U'\in A,V'\in B)=\pi}\ \inf_{\omega'\in\Omega'} \min[X_\omega(U'(\omega')),Y_\omega(V'(\omega'))]$$

定理 5.3.4 如定理 5.3.1 那样定义 $r'_A(\omega)$ 和 $r''_A(\omega)$，且令

$$S'_B(\omega)=\sup_{y\in B}\boldsymbol{Y}_\omega(y),\quad S''_B(\omega)=\sup_{y\in B^c}\boldsymbol{Y}_\omega(y)$$

则

$$I_\omega^{X\in A,Y\in B}(\pi)=\begin{cases}\max[r''_A(\omega),S''_B(\omega)], & \pi=0.\\ \min[r'_A(\omega),S'_B(\omega),\max[r''_A(\omega),S''_B(\omega)]], & 0<\pi<1,\\ \min[r'_A(\omega),S'_B(\omega)], & \pi=1.\end{cases}$$

证明类似于定理 5.3.2 的证明。

定理 5.3.5 $EI^{X\in A,Y\in B}=EI^{X\in A}I^{Y\in B}$.

证明是显然的。

定理 5.3.6 若 $\boldsymbol{X}$，$\boldsymbol{Y}$ 独立，则

$$P_r(\boldsymbol{X}\in A,\ \boldsymbol{Y}\in B)=P_r(\boldsymbol{X}\in A)P_r(\boldsymbol{Y}\in B).$$

证 因 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 独立，所以 $I^{X\in A}$和 $I^{Y\in B}$也独立。从定理5.3.5和定理5.3.1得到。

当 $\boldsymbol{X}$ 是离散 KF随机变量时，可得计算$\boldsymbol{F}$概率$\boldsymbol{Pr}(\boldsymbol{X}\in\boldsymbol{A})$的规则如下：设 $\boldsymbol{X}=\{(p_i,X^i),\ i=1,2,\cdots\}$

步骤1 对每个 i 确定

$$r_i'=\sup_{x\in A}X^i(x),\quad r_i''=\sup_{x\in A^c}X^i(x),\quad i=1,2,\cdots$$

步骤2 由 $P_r(\boldsymbol{X}\in A)$、$\min(r_i',r_i'')$、$\mu_0=1$ 及 $\mu_K\neq 1$，$(K=1,2,\cdots)$的假定，确定水平 μ_k，$k=0,1,2,\cdots$，并整理μ_K成如下顺序，即

$$1=\mu_0>\mu_1>\cdots>\mu_n>\cdots.$$

步骤3 对每个 $k=0,1,2,\cdots$和 $q\in[0,1]$ 确定

$$\chi_K(q)=\begin{cases}1, & \displaystyle\sum_{i:\,r_i''<\mu_k}p_i\leqslant q\leqslant 1-\sum_{i:\,r_i'<\mu_k}p_i\\ 0. & \text{其它}\end{cases}$$

步骤4 对于 $q\in[0,1]$确定

$$(P_r(\boldsymbol{X}\in A))(q)=\sup_K\chi_K(q)\mu_K.$$

例5.3.1 就意见征集的例子来考虑。已知$p_1=0.4$，$p_2=0.5$，$p_3=0.1$，$p_K=0(K\geqslant 4)$，且"很暖"，"暖"及"不评价"的隶属函数分别为 X^1，X^2及 X^3。现在来求 $P_r(\boldsymbol{X}\geqslant z)$，$z\in R$，于是 $A=[z,+\infty)$。对于给定的 z 和 q，$P_r(\boldsymbol{X}\geqslant z)$是反映赞成温度超过 z℃意见的一个分数 q。利用上述规则，对于给定的 z 值可很容易地确定 $P(\boldsymbol{X}\geqslant z)$。事实上，若设 $z=25$℃。

步骤 1 $r_1' = \sup_{x\in(25,+\infty)} X^1(x) = 1$,

$r_1'' = \sup_{x\in(-\infty,25)} X^1(x) = 0$,

$r_2' = 1$, $r_2'' = 1$; $r_3' = 1$, $r_3'' = 1$.

步骤 2 $\min(r_i', r_i'') = 0$, $i = 1,2,3$,取$\mu_0 = 1$,$\mu_2 = 0.5$, $\mu_3 = 0.4$, $\mu_4 = 0.3$ 等。

步骤 3 对 $K = 0,1,2,3$, 有

$$\sum_{i:r_i''<\mu_k} p_i = p_1 = 0.4, \qquad 1-\sum_{i:r_i'<\mu_k} p_i = 1-0=1.$$

所以当 $q\in[0,1]$时, 有

$$\chi_0(q) = \chi_1(q) = \chi_2(q) = \chi_3(q) = \begin{cases} 1, & 0.4\leqslant q<1, \\ 0, & 其它, \end{cases}$$

步骤 4 得 $(P_r(\boldsymbol{X}\geqslant 25))(q) = \sup_k \chi_k(q)\mu_k$, 故对于 $q\in[0,1]$, 得 $P(\boldsymbol{X}\geqslant 25)$ = "至少 0.4"的结论。如此对于不同的 Z 值. 类似上面计算可以得出 $P(\boldsymbol{X}\geqslant Z)$之值。如 z = 30℃, 有 $P(\boldsymbol{X}\geqslant 30)$ = "不比 0.5 大"。而当 Z = 25,26,27,28,29 和 30℃ 时, $P(\boldsymbol{X}\geqslant Z)$从"至少为 0.4"逐渐过渡到"不比 0.5 大"。

第六章

F 统计

有些时候，人参与了某些随机现象，使得其中既含有不分明性又包含着随机性。例如在人机系统、管理系统、经济系统及社会系统中的某些情况就是如此。对这些现象如何进行统计分析，有着十分重要的现实意义。目前这方面的研究进展，远不能满足实际的需要，但是已有相当的基础。

我们这一章主要用来介绍不分明观念与统计相结合所产生的一些新结果。它们是 F 统计法，F 正交实验设计及 F 相关分析。由于与统计密切关联，故统一在 F 统计这个标题之内。而 F 聚类分析另辟一章介绍。

§1 F 统计法

F 统计法是用来确定 F 集合的隶属函数的一种方法，其数学模型本身不具有随机性，但其思想方法非常类似于统计模型，从而可看作是 F 集思想与统计方法的结合。

6.1.1 统计方法的回顾

在经典集合论基础上建立的统计理论中，随机试验是一个最基本的概念，由它引出频率的定义，在引进经验分布概念之后开展了一系列的统计工作。随机试验涉及以下四个要素：

(1) 样本空间 Ω；

(2) 事件 $A\subset\Omega$；A 是 Ω 上的一个确切集合；

(3) 样本点 ω，由它确定事件 A；

(4) 条件 S。它是对变元 ω 活动的一种限制。

在此我们所关心的随机性，其产生的根本原因在于条件 S 的不充分，它未将 ω 完全限制死，仍有活动的余地，而且

$$S \cap A \neq \phi \neq S \cap \overline{A}。$$

当 $\omega \in S \cap A$ 时，事件 A 发生；当 $\omega \in S \cap \overline{A}$ 时，事件 A 不发生。故 A 是条件 S 之下的随机事件。

随机试验的目的：用确定性手段去研究随机性。

在随机试验中的每次试验中，A 是固定的，ω 是变化的。A 发生与否，必须是明确的。

作 n 次试验，定义 A 发生的频率为 ω 发生的次数与 n 的比值。随着 n 的增大，频率通常会呈现出稳定性，频率稳定所在的那个数，就称为事件 A 在 S 下的概率。

6.1.2 不分明统计

什么是不分明统计呢？它和概率统计有着什么本质区别呢？我们设想作一个试验，考虑在人的集合（论域 U）中，张松（u_0）是否是高个子（A^*）。显然，$u_0 \in U$。而 A^* 是 U 的一个运动着的，边界可变的集。这是因为“高个子”这个概念是随着不同的条件、场合和观点而改变的。每次试验也是清楚的可理解的，即让不同的人来评论张松是不是属于高个子这个集合 A^*。当然有时会得到“张松属于高个子”，即 $u_0 \in A^*$；但有时也会得到“张松不是高个子”，即 $u_0 \notin A^*$。于是，作 n 次试验后，总可以得到 $u_0 \in A^*$ 在 n 次试验中发生的次数，从而它与 n 的比就可用来定义 u_0 对 A^* 的隶属频率：

$$u_0\text{对 }A^*\text{的隶属频率} = \frac{u_0 \in A^*\text{的次数}}{n}。$$

许多实践证明[2]，随着 n 的增大，隶属频率也会呈现稳定性，我们称之为隶属频率的稳定性。频率稳定所在的那个数，叫做 u_0对 A^*的隶属度。

将隶属度不为 0 的那些元素 u_0和 u_0的隶属度 μ_0 搭配起来组成"单点"集，即$\{(\mu_0, u_0)\}$，即得"高个子"这一F 集合.

这里的试验，也有四个要素：

(1) 论域U，它是确定的集合；

(2) 元素 u_0，它是 U 的一个固定元素；

(3) $A^* \subset U$，它是一个运动着的，边界可变的普通集合. A^*联系于一个不分明集合 $\underset{\sim}{A}$(相应于不分明概念 α)，A^*的每一次固定化，都是对 α 所作出的一个确定划分，它表示 α 的一个近似的外延，A^*也可看成 $\underset{\sim}{A}$ 的一次实现，$\underset{\sim}{A}$ 的疆域控制着 A^*的变化；

(4) 条件 S，它联系对概念 α 所进行的划分过程的全部客观或心理的因素，制约着 A^*的运动。

由于 S 对划分过程没有限制死，A^*可以变异，它可以覆盖 u_0，也可以不覆盖 u_0，致使 u_0对 A^*的隶属关系是不确定的。

在每次试验时，A^*是取定的普通集合。而在各次试验中，u_0是固定的，A^*在变化.

称以上方法为不分明统计法，简称 F 统计法。它是用确定的手段研究不分明性。

例 6.1.1 设$U = (0, 100)$(单位：岁)，$\underset{\sim}{A}$ 是"青年人"在 U 上的 F 集，选取 $u_0 = 27$。试用 F 统计试验，来确定 u_0 对 $\underset{\sim}{A}$ 的隶属程度。

张南纶，蔡训武[82], [2]在武汉建材学院作抽样试验，

表 1　　　　　　　　武汉建材学院抽样情况

样本总数 $n=129$（无作废数据），单位是岁.

全部数据如下：

19-25	17-30	17-28	18-25	16-35	14-25	18-30	18-35	18-35	16-25
15-30	18-35	17-30	18-25	18-25	18-35	20-30	18-30	16-30	20-35
18-30	18-30	17-25	19-30	15-28	16-28	18-30	18-30	16-30	18-35
18-25	18-25	16-28	18-30	16-30	16-28	18-35	18-35	17-27	16-28
15-28	16-30	19-28	15-30	15-26	17-25	15-36	18-30	17-30	18-35
16-35	15-25	15-25	18-28	16-30	15-28	18-35	18-30	17-28	18-35
15-28	18-30	15-25	15-25	18-30	16-24	15-25	16-32	15-27	18-35
16-25	18-28	16-28	18-30	18-35	18-30	18-30	17-30	18-30	18-35
16-30	18-35	17-25	15-30	18-25	17-30	14-25	18-26	18-29	18-35
18-28	18-30	18-25	16-35	17-29	18-25	17-30	16-28	18-30	16-28
15-30	15-35	15-30	20-30	20-30	16-25	17-30	15-30	18-30	16-30
18-28	18-35	16-30	15-30	18-35	18-35	18-30	17-30	16-35	17-30
15-25	18-35	15-30	15-25	15-30	18-20	17-25	18-29	18-28	

选择 129 人，在他们独自认真考虑了“青年人”的含义之后，报出了他们认为“青年人的最适宜的年限，具体数据如 表 1.

表 1 所列数据，每一个表示答复者认为“青年人”所包括的年令，如 19-25，指“青年人”范围是 19 岁至 25 岁。可见，每一个答复相当于一次试验，即给定一个具体的 A^*。

将表 1 所列数据进行分组，并计算出隶属频率。因为试验共进行了 129 次，故最高隶属频数为 129，即 $u_0 \in A^*$的次数最高为 129。表 1 中最小岁数为 14 岁，最大岁数 为 36 岁.

表 2　　分组计算相对频数

序号	分　　组	频数	相对频数	序号	分　　组	频数	相对频数
1	13.5～14.5	2	0.0155	13	25.5～26.5	103	0.7984
2	14.5～15.5	27	0.2093	14	26.5～27.5	101	0.7829
3	15.5～16.5	51	0.3953	15	27.5～28.5	99	0.7674
4	16.5～17.5	67	0.5194	16	28.5～29.5	80	0.6202
5	17.5～18.5	124	0.9612	17	29.5～30.5	77	0.5969
6	18.5～19.5	125	0.9090	18	30.5～31.5	27	0.2093
7	19.5～20.5	129	1	19	31.5～32.5	27	0.2093
8	20.5～21.5	129	1	20	32.5～33.5	26	0.2016
9	21.5～22.5	129	1	21	33.5～34.5	26	0.2016
10	22.5～23.5	129	1	22	34.5～35.5	26	0.2016
11	23.5～24.5	129	1	23	35.5～36.5	1	0.0087
12	24.5～25.5	128	0.9922			$\sum$	13.6589

采用组距为 1 可分为 23 组：从 13.5 岁开始至 36.5 岁止。

表 1 中第一个答复为 19-25，共包括 7 个岁数。从而使第 6 组至第 12 组各出现一次。而 14 岁这个年令，仅出现两数，故频数为 2。于是相对频数为 2/129 = 0.0155。现将129个数据分组计算相对频数，结果列于表 2。

以各组的组中值为代表，从而得其隶属频率。如第14组的组中值为 27 岁，于是得 27 岁隶属于“年青人”的隶属度为 0.7829。如此求出各组的组中值．结合相应的隶属度即得相应的隶属函数曲线及相应的“年青人”这一F 集。

以上介绍的 F 统计法是所谓的二相 F 统计，二相指的是：

$$P_2 = \{\underset{\sim}{A}, \underset{\sim}{A}^c\},$$

每进行一次不分明试验，都确定了一个映射：

$$e: U \to P_2$$

它是对 U 的一次划分。二相 F 统计等于是两个相反的 F 概念在 U 中进行“竞选”的统计。

显然，存在着多相 F统计，关于它我们就从略了（见[2]）。

§2 F 正交试验设计

在数理统计中，正交试验设计法占有一定的比重，在工农业生产方面很有用。但当涉及不分明因素时，就不能应用。能否将该方法应用于 F 领域，是值得考虑的一个问题．为此目的，这节介绍这方面的一些研究，虽然有待于进一步深化和开垦，但仍是一个好的开端。

为了帮助不了解正交试验设计的读者，我们先回顾一下这方面的内容，然后介绍 F 正交试验设计。

6.2.1 正交试验设计

在这一节，我们首先简单谈一下正交试验设计。

我们遇到的实际问题，常常是比较复杂的，包括有多种因素，各个因素又有不同的状态，它们相互交织在一起。为了寻求影响指标的最合适的生产条件，就要对各种因素以及各个因素的不同状态进行试验。

例 6.2.1 某农药厂生产某种农药，根据生产经验，发现影响农药收率的因素有 4 种，每种因素都有两种状态，具体如下：

反应温度(A)：60℃(A_1)，80℃(A_2)，

反应时间(B)：2.5 小时(B_1)，3.5 小时(B_2)，

配比(C)：　1.1/1(C_1)，1.2/1(C_2)

真空度(D)：　500 毫米汞柱(D_1)，600 毫米汞柱(D_2)。

问题是找出各因素对指标(收率)的影响规律，即寻求出哪个因素是主要的，哪个是次要的？哪些因素只起单独作用，哪些因素除了各自的单独作用外，它们之间还产生综合效果？这种综合效果有多大？对指标的影响，综合效果是主要的，还是因素的单独作用是主要的？另外，选出各因素的一个状态来组成比较合适的生产条件。

要解决这些问题，就必须进行实验或试验。为了便于讨论，以后通称影响指标的因素为因子，用大写字母 A，B，…表示，每个因子可能处的状态称为水平，用字母加足标表示。如 A_1、A_2，…表示 A 因子的第一、第二、…水平等。有时，亦将因子用大写字母加单足标表示， 如 A_1，A_2，…表示第一个因子，第二个因子，…等，而可能水平用相应因子符号加第二个足标表示，如 A_{11}，A_{12}，… 等表示因子

A_1 的第一个水平，A_1 的第二个水平，…等。

在上述例子中，有四个因子，每个因子有两个水平。因此，上述问题称为四个两水平因子的试验问题。对于此问题要寻求圆满的解决，就必须进行 $2^4=16$ 次试验结果的处理。如果因子和水平的数目比较多时，要解决这样一个问题就要作更多次试验。因此产生一个问题，是否可以只作其中一部分试验就能解决上面所提问题呢？要作到这一点，就必须合理的安排试验。这种合理安排的方案就是正交表，用正交表来安排试验的方法称为正交试验设计法。

最简单的正交表是 $L_4(2^3)$。其中 L 代表正交表，L 下角的数字 4 表示要作四次试验，即表中有四横行（简称为行）；括号内的数字 2，表示表的主要部分只有 2 种数字，即因子有两种水平 1 和 2，称之为 1 水平与 2 水平；括号内的指数 3 表示有 3 纵列（简称列），即最多允许安排的因子个数是 3 个。具体如下：

列号 试验号	1	2	3
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

一般来说，正交表 $Ln(K^i)$ 表示：

正交表 ——→ $Ln(K^i)$

需作试验数 —— ↑（n）

水平数 —— ↑（K）

最多安排因子数 —— ↓（i）

根据问题中因子的个数及水平的个数，选择相应的正交表(各种正交表见[5])．根据表的规定，制订试验方案．如上述例子的四个两水平试验不能用 $L_4(2^3)$ 来安排，因为 $L_4(2^3)$ 最多只能安排 3 个因子两个水平的试验，所以我们选 $L_8(2^7)$ 来制订方案。$L_8(2^7)$ 为

试验号 \ 列号	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

由该表来安排试验：将因子放于列号内，我们选取 1，2，4，7 分别为 A，B，C，D 四因子．并将各因子所处列中数字 1 和 2 的位置分别换上该因子的 1 水平和 2 水平，得计划表如下：

下表中收率表示试验结果．如第五次试验是 A_2，B_1，C_1、D_2 搭配(A_2，B_1，C_1，D_2)后进行的试验，所得收率为 91%。其余的含义相同。

有了试验结果，即可进行分析处理数据。通常最常用的方法是直观分析法与方差分析法。由于我们的目的是介绍正交试验设计的 F 分析法，故此从略。读者可在数理统计的书中查到。

因子 \ 试验号	A	B	C	D	收率(%)
1	60	2.5	1.1/1	500	86
2	60	2.5	1.2/1	600	95
3	60	3.5	1.1/1	600	91
4	60	3.5	1.2/1	500	94
5	80	2.5	1.1/1	600	91
6	80	2.5	1.2/1	500	95
7	80	3.5	1.1/1	500	83
8	80	3.5	1.2/1	600	88

6.2.2 *F* 正交试验设计

我们利用一个具体问题来介绍不分明 $L_9(3^4)$正交试验设计(见[33])。

设论域为 $U=\{x_1, x_2, x_3\}$。它由三个子论域 x_1, x_2, x_3 构成，而且分别作为三个因素 A_1(季节)，A_2(整地方式)及 A_3(树种)的论域。每个因素取三个水平，即

因素 \ 序号	A_1	A_2	A_3	成活率(%)
1	A_{11}	A_{21}	A_{31}	0.88
2	A_{11}	A_{22}	A_{32}	0.87
3	A_{11}	A_{23}	A_{33}	0.87
4	A_{12}	A_{21}	A_{32}	0.56
5	A_{12}	A_{22}	A_{33}	0.65
6	A_{12}	A_{23}	A_{31}	0.86
7	A_{13}	A_{21}	A_{33}	0.86
8	A_{13}	A_{22}	A_{31}	0.80
9	A_{13}	A_{23}	A_{32}	0.86

A_1：早春 A_{11}，雨季 A_{12}，晚秋 A_{13}，

A_2：大鱼鳞坑 A_{21}，反坡梯田 A_{22}，水平沟 A_{23}，

A_3：侧柏 A_{31}，油松 A_{32}，刺槐 A_{33}。

用 $L_9(3^4)$来安排试验，所得到的成活率整理如下：

K_1	2.62	2.30	2.54
K_2	2.07	2.32	2.29
K_3	2.52	2.59	2.38
K'_1	0.873	0.767	0.850
K'_2	0.690	0.773	0.763
K'_3	0.840	0.863	0.793

我们记因素 $A_i(i=1,2,3)$在 $j(j=1,2,3)$ 水平上 A_{ij} 的平均成活率为 $a_{ij}(0\leqslant a_{ij}\leqslant 1)$，认为是 A_i 对其在 j 水平 $A_{ij}\in x_i$ 处"成活率高"的隶属度对造林成活率提高的影响度，因此 A_i 相对于各水平的平均成活率(a_{i1},a_{i2},a_{i3})也是定义在x_i上的 F 集 $\underset{\sim}{A}_i$，而(a_{i1},a_{i2},a_{i3})在 x_i 上的表现，是由于人不能对全部因素的一切水平都进行考察，就把它压缩在低维空间 x_i 上进行考察所得到的结果.这种表现，实际上蕴含着不分明性。因此

$$\underset{\sim}{A}_1=(0.873,\ 0.690,\ 0.840),$$

$$\underset{\sim}{A}_2=(0.767,\ 0.773,\ 0.863),$$

$$\underset{\sim}{A}_3=(0.850,\ 0.763,\ 0.793).$$

分别为 x_1，x_2，x_3 上的 F 集。

因为 F 集 $\underset{\sim}{A}_1$ 的高度 $\text{hgt}\ \underset{\sim}{A}_1=0.873$，而 $\text{hgt}\ \underset{\sim}{A}_2=0.863$，$\text{hgt}\ \underset{\sim}{A}_3=0.850$，故

$$\text{hgt}\ \underset{\sim}{A}_1>\text{hgt}\ \underset{\sim}{A}_2>\text{hgt}\ \underset{\sim}{A}_3$$

因此各因素对造林成活率的影响，以季节为最大，其次是整

地方式，再次是树种。

为了分析正交试验设计中任意两个因素的各种水平搭配对造林成活率的影响，对 F 向量 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，$\underset{\sim}{A}_3$ 中每两个之间的笛卡尔乘积定义为 $\underset{\sim}{R}_i(i=1,2,3)$：$\underset{\sim}{R}=(r^i_{jk})$，$r^i_{jk}=a_{ij}\wedge a_{i+1,k}$，其中 $\underset{\sim}{A}_i=(a_{i1},a_{i2},a_{i3})(i=1,2,3)$。$\underset{\sim}{R}_i$ 是论域 $X_i\times X_{i+1}$ 上的 F 关系，r^i_{jk} 表示因素 $\underset{\sim}{A}_i$ 在 j 水平处与因素 $\underset{\sim}{A}_{i+1}$ 在 k 水平处搭配对造林成活率的联合影响。

如

$$\underset{\sim}{A}_1\times\underset{\sim}{A}_2\triangleq\begin{pmatrix}0.873\wedge0.767, & 0.873\wedge0.773, & 0.873\wedge0.863\\ 0.690\wedge0.767, & 0.690\wedge0.773, & 0.690\wedge0.863\\ 0.840\wedge0.767, & 0.840\wedge0.773, & 0.840\wedge0.863\end{pmatrix}$$

$$=\begin{pmatrix}0.767, & 0.773, & 0.863\\ 0.690, & 0.690, & 0.690\\ 0.767, & 0.773, & 0.840\end{pmatrix}=\underset{\sim}{R}_1.$$

为了分析当一因素固定在某一水平时，它与其它因素不同水平搭配时对造林成活率的影响。引入不分明关系 $\underset{\sim}{R}_i$ 在论域 X_k 中的投影，记为 $\underset{\sim}{R}_{iX_k}$，它由隶属函数

$$\mu_{\underset{\sim}{R}_{iX_k}}(x_j)=\bigvee_{x_j\in X_k}\mu_{\underset{\sim}{R}_i}(x_1,\ x_2).$$

所确定。投影表示不分明关系矩阵的行峰值和列峰值。如 $\underset{\sim}{R}_{1X_1}$，其隶属函数

$$\mu_{\underset{\sim}{R}_1X_1}(x_1)=(0.863,\ 0.690,\ 0.840)',$$

$$\mu_{\underset{\sim}{R}_1X_2}(x_1)=(0.767,\ 0.773,\ 0.863).$$

这里 $\underset{\sim}{R}_1$ 是造林季节与整地方式对造林成活率的联合影

响。从而投影可用来分析当一因素固定在某一水平时它与其它各因素不同水平搭配时对造林成活率的影响。从上看出，水平搭配以 $A_{11}A_{23}$ 为最好，对成活率的影响最有利，可达到 0.863。

同样办法可分析出搭配 $A_{23}A_{31}$ 可达 0.850，$A_{31}A_{11}$ 可达 0.850。

为了分析所有因素的各种水平搭配对造林成活率的影响，定义 $\underset{\sim}{A}_1, \underset{\sim}{A}_2, \underset{\sim}{A}_3$ 之间的笛卡尔乘积：

$$\underset{\sim}{A}_1 \times \underset{\sim}{A}_2 \times \underset{\sim}{A}_3 \triangleq \int_{x_1 \times x_2 \times x_3} \frac{\min(a_{1i}, a_{2i}, a_{3i})}{(A_{1i}, A_{2i}, A_{3i})}$$

例如，$(A_{11},\ A_{22},\ A_{31})$关于 F 集 $\underset{\sim}{A}_1 \times \underset{\sim}{A}_2 \times \underset{\sim}{A}_3$ 的隶属度是 $0.873 \wedge 0.773 \wedge 0.850 = 0.773$，它表示早春季节、反坡梯田和侧柏的搭配对造林成活率的联合影响是 0.773。容易看出，搭配(A_{11}, A_{23}, A_{31})在各种可能组合中对造林成活率的影响最大，其值为 $0.873 \wedge 0.863 \wedge 0.850 = 0.850$。

综上所述，在不考虑交互作用的情况下，三个因素的水平搭配为早春、水平沟和侧柏时其成活率最高。

作者在 1982 年及 1983 年在山区作了实验，效果与上述结果基本一致.说明他的试验方案是行之有效的.而且说明应用正交试验设计的 F 分析法，不仅符合客观实际，同时说明了该方法的可行性。

§3 F 线性回归分析

回归分析是数理统计中一个重要的内容，在实际方面也有多种应用。由于它不能使用于不精确的数据，从而就必需将其扩展至不分明领域。这一节就是用来介绍这一主题。

为了方便，我们首先介绍简单的二元线性回归，然后再介绍 F线性回归[34]。

6.3.1 二元线性回归

假如自变量 x 与因变量 y 的实测值为

$$(x_i, y_i), \quad i=1, 2, \cdots, n.$$

在坐标纸上这 n 组值就是 n 个点。现在的问题是去寻求 x 与 y 之间的近似表达式。若这 n 个点似乎位于一直线附近，就可认为 y 与 x 之间有线性关系，不妨设其为 $y=ax+b$。因此 x_i 与 y_i 有关系式

$$y_i = a + bx_i + \varepsilon_i, \quad i=1,2,\cdots,n. \tag{6.3.1}$$

其中 x_i，y_i 为已知的，a，b，ε_i 是未知的。怎样利用关系式(6.3.1)来求出 a，b 呢？这可利用最小二乘法。

将(6.3.1)式改写为

$$\varepsilon_i = y_i - a - bx_i, \quad i=1,2,\cdots,n \tag{6.3.2}$$

它表示用 $a+bx_i$ 来代替 y_i 所产生的误差。全部误差的平方和是

$$\sum_{i=1}^{n}\varepsilon_i^2 = \sum_{i=1}^{n}(y_i - a - bx_i)^2, \tag{6.3.3}$$

记为 Q。我们要求的 a，b 应该使(6.3.1)式符合得好，也就是误差项 ε_i 要小。因此 Q 就是反映符合好坏的一个量。最小二乘法告诉我们应选 a，b 使 Q 达到最小值。

我们对(6.3.3)式求 Q 对a，b 的偏导数，并令其为零，得

$$\begin{cases} \sum_{i=1}^{n}(y_i - a - bx_i) = 0 \\ \sum_{i=1}^{n}(y_i - a - bx_i)x_i = 0 \end{cases} \tag{6.3.4}$$

将其改写为

$$na + \left(\sum_{i=1}^{n} x_i\right)a_i = \sum_{i=1}^{n} y_i$$

$$\left(\sum_{i=1}^{n} x_i a\right) + \left(\sum_{i=1}^{n} x_i^2\right)b = \sum_{i=1}^{n} x_i y_i$$

对 a，b 求解。将一式两端除以 n，得 $a = \bar{y} - b\bar{x}$，代入第二式就得 b。我们将其解记为 $\hat{a}$，$\hat{b}$，即有

$$\begin{cases} \hat{a} = \bar{y} - b\bar{x} \\ \hat{b} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2. \end{cases} \tag{6.3.5}$$

这样我们求出了 a，b 这两个值。a 称为回归常数。b 称为回归系数。$\hat{a}$，$\hat{b}$ 分别为 a，b 的最小二乘估计值。于是

$$\hat{y} = \hat{a} + \hat{b}x \tag{6.3.6}$$

称为 y 的估计值。也是从 $(x_i,\ y_i)$，$i = 1,\cdots,n$ 资料得到的回归方程。不难验证，如此 $\hat{a}$，$\hat{b}$ 的确使 Q 达到最小。

以上介绍的方法和近似计算中的曲线拟合是一致的。而在数理统计中则还需要考虑以下三个问题：

(1) 回归方程是否有意义？

(2) 若有意义，x 与 y 相关到什么程度？

(3) 若有意义，如何用 x 来预测 y？

为了解决这些问题，需要引进相关系数的概念：两个随机变量 ξ 与 ζ 的相关系数 ρ 定义为

$$\rho=\mathrm{cov}(\xi,\eta)/\sqrt{D\xi\cdot D\eta} \tag{6.3.7}$$

其中 $\mathrm{cov}(\xi,\eta)$ 为 ξ 与 η 的协方差，而 $D\xi$ 及 $D\eta$ 分别为 ξ 与 η 的方差。当 $(\xi,\ \eta)$ 的样本值为 $(x_i,\ y_i)$，$i=1,\cdots,n$ 时，则有样本相关系数 R：

$$R=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2\cdot\sum_{i=1}^{n}(x_i-\bar{x}^2)}}$$

其中 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}\bar{y}$。通常把 R 记为 $\hat{\rho}$，即 ρ 的估计值。

从实际观测值（即样本）$(x_i,\ y_i)$ 得到相关系数 R。我们由 R 的大小即可判定 x 与 y 之间的相关程度，而且当 R 很小时，就认为 x 与 y 之间无线性关系，即所配回归无意义，否则，就认为所得回归方程有意义。至于第三个问题，对于学过数理统计的读者来说，是明显的，这里从略了。

从以上介绍可以看出，回归分析中涉及这样的一些概念：误差、距离、相关系数。它们在F 回归分析中是必须加以模糊化的一些概念。

6.3.2 F 相关分析

当样本点 (x_i,y_i) 不是 R^2 中一些精确的数据，而是一些可能取值的范围，并且这些范围的边界不清晰时，上述方法就不能被用来对样本点进行相关分析。因此有必要将上述

方法予以推广。这就是此处的不分明相关分析，简称为 F 相关分析。

根据 F 集论的观点，如此的样本点可用 R^2 中的 F 集 $\underset{\sim}{A}$ 来描述，并称为 F 样本点. 设我们已测得 n 个 F 样本点 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，…，$\underset{\sim}{A}_n$。问题是如何对它们进行线性拟合。

为了解决此问题，我们先引进 F 距离。

定义 6.3.1 设 $\underset{\sim}{A}\in\mathscr{F}(R^2)$，精确直线 l 的方程为 $y=ax+b$，则 $\underset{\sim}{A}$ 到直线 l 的不分明距离（F 距离）定义为 $\underset{\sim}{D}=f(\underset{\sim}{A})$，$\underset{\sim}{D}\in\mathscr{F}(R^2)$。$\underset{\sim}{D}$ 的隶属函数为

$$\mu_{\underset{\sim}{D}}(d)=\sup_{\substack{x,y\\ d=f(x,y)}}\mu_{\underset{\sim}{A}}(x,\ y).$$

其中 $d=f(x,y)$ 是平面上点 $A(x,y)$ 到直线 l 的距离。它可以取为

$$f(x,\ y)=|y-ax-b|.$$

为了讨论 $\underset{\sim}{D}$ 的性质，我们首先讨论一个引理。

引理 1 设 U 和 V 是两个论域，$f: U\to V$，$\underset{\sim}{M}\in\mathscr{F}(U)$，$f(\underset{\sim}{M})\in\mathscr{F}(V)$，则 $\forall d\in(0,1)$，

$$[f(\underset{\sim}{M})]_\alpha=f(M_\alpha)$$

的必充条件为：$\forall v\in V$，存在 u^* 使得

$$\mu_{f(\underset{\sim}{M})}(v)=\mu_{\underset{\sim}{M}}(u^*).$$

性质 1 设 $\underset{\sim}{A}\in\mathscr{F}(R^2)$，$\underset{\sim}{A}$ 有连续隶属函数 $\mu_{\underset{\sim}{A}}(x,\ y)$，若 $\forall\alpha\in(0,\ 1)$，A_α 有界，则 F 距离 $\underset{\sim}{D}$ 是 R^+ 上的凸 F 集。

证 由距离函数 $f(x,\ y)$ 连续和 $\underset{\sim}{A}$ 隶属函数 $\mu_{\underset{\sim}{A}}(x,\ y)$ 连续，则 $\forall d^*\in R^+$，存在 $(x^*,\ y^*)\in R^2$，使得

$$\mu_{\underset{\sim}{D}}(d^*)=\mu_{\underset{\sim}{A}}(x^*,\ y^*)$$

由引理 1，$\forall\alpha\in(0,1)$，有 $D_\alpha=f(A_\alpha)$。

由 $\mu_{\underset{\sim}{A}}(x,y)$的连续和 A_α 的有界，则 A_α 是闭集，并由上式可知 D_α 是连续闭区间。所以 $\underset{\sim}{D}$是 R^+ 上的凸 F 集。

性质 2 在性质 1 的条件下，若存在$(x^*,y^*)\in R^2$，使得$\mu_{\underset{\sim}{A}}(x^*,y^*)=1$，则 $\underset{\sim}{D}$是 R^2 上的 F 数。

根据 F 数的定义，结论是显然的。

在实际问题中，对实测量可以有一个允许的误差范围，并且这个范围可以有弹性，因此可以定义一个不分明误差标准集，简称为 F 误差标准集。

定义 6.3.2 设正实数 c 是容许误差，正实数 $d>c$ 是最大容许误差，则 R^+ 上的 F 集 $\underset{\sim}{Q}$ 称为 F 误差标准集，$\underset{\sim}{Q}$ 的隶属函数规定如下：

$$\mu_{\underset{\sim}{Q}}(t)=\begin{cases}1 & 0\leqslant t\leqslant c,\\ \dfrac{d-t}{d-c} & c<t<d,\\ 0 & t\geqslant d.\end{cases}$$

显见 $\underset{\sim}{Q}$ 是 R^+ 上的 F 数，且对某个 $t^*\in R^+$，$\mu_{\underset{\sim}{Q}}(t^*)$值越接近于 1，表示对误差要求的满意程度越高。

设 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，…，$\underset{\sim}{A}_n$ 是在 R^2 中所测得的 F 样本点，其拟合直线设为 $l: y=ax+b$。R^+ 上的 F 集 $\underset{\sim}{D}_1$，$\underset{\sim}{D}_2$，…，$\underset{\sim}{D}_n$ 分别是这 n 个 F 样本点到直线 l 的 F 距离，则我们希望直线 l 的选取使得所有的 F 距离 $\underset{\sim}{D}_1$，$\underset{\sim}{D}_2$，…，$\underset{\sim}{D}_n$ 都"较小"，并"较小"到所要求的范围。为了对这种"较小"进行刻划，我们先引进 R 上两个 F 集大小比较的数量指标。

定义 6.3.3 设 $\underset{\sim}{M}$，$\underset{\sim}{N}\in\mathscr{F}(R)$，由扩张原理，$F$ 集$\underset{\sim}{M}$

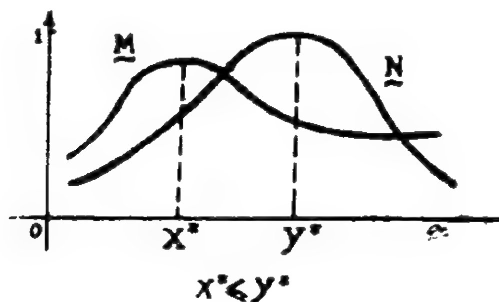
"小于"F 集$\underset{\sim}{N}$ 的数量指标规定为

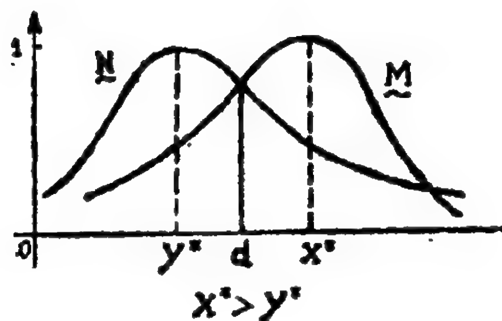
$$\nu(\underset{\sim}{M}\leqslant\underset{\sim}{N})=\sup_{\substack{x,y\\x\leqslant y}}\min(\mu_{\underset{\sim}{M}}(x),\ \mu_{\underset{\sim}{N}}(y)).$$

设$\underset{\sim}{M}$ 和$\underset{\sim}{N}$ 都是凸 F 集，若 $\sup\limits_{x}\mu_M(x)=\mu_M(x^*)$，$\sup\limits_{y}\mu_{\underset{\sim}{N}}(y)=\mu_{\underset{\sim}{N}}(y^*)$，则

i) $x^*\leqslant y^*$ 时，$\nu(\underset{\sim}{M}\leqslant\underset{\sim}{N})=\min\{\mu_{\underset{\sim}{M}}(x^*),\ \mu_{\underset{\sim}{N}}(y^*)\}$，

ii) $x^*>y^*$ 时，$\nu(\underset{\sim}{M}\leqslant\underset{\sim}{N})=\mathrm{hgt}(\underset{\sim}{M}\cap\underset{\sim}{N})$.





i) $x^*\leqslant y^*$　　　　ii) $x^*>y^*$

图 6.3.1　$\nu(\underset{\sim}{M}\leqslant\underset{\sim}{N})$的示意图

有了这些概念之后，可见在给定一个 F误差标准 Q的情况下，任何一个 F 样本点 A_i 与所拟合直线 l : $y=ax+b$ 的

F 距离$\underset{\sim}{D}_i$，可用

$$\nu_i = \nu(\underset{\sim}{D}_i \leqslant \underset{\sim}{Q})$$

来表示$\underset{\sim}{D}_i$是"较小"到满足 F 误差标准 $\underset{\sim}{Q}$ 的可信度。

利用可信度，可以定义相关系数的概念。

定义 6.3.4 设 n 个 F 样本点 $\underset{\sim}{A}_1, \cdots, \underset{\sim}{A}_n \in \mathscr{F}(R^2)$，它们到直线 l：$y=ax+b$ 的 F 距离分别为 $\underset{\sim}{D}_1$，…，$\underset{\sim}{D}_n$。$\underset{\sim}{Q}$ 是给定的 F 误差标准集，则 n 个 F 样本点 $\underset{\sim}{A}_1, \underset{\sim}{A}_2, \cdots, \underset{\sim}{A}_n$ 到直线 l 的 F 相关系数定义为

$$R = \sum_{i=1}^{n} \lambda_i \nu_i = \sum_{i=1}^{n} \lambda_i \nu(\underset{\sim}{D}_i \leqslant \underset{\sim}{Q}).$$

其中 $\lambda_i \geqslant 0$，$\sum_{i=1}^{n} \lambda_i = 1$，$\lambda_i$ 表示 F 样本点 $\underset{\sim}{A}_i$ 重要性的权数。

显见 $0 \leqslant R \leqslant 1$。

当在 F 误差标准 $\underset{\sim}{Q}$ 给定的前提下，R 的值越大，则用直线 l：$y=ax+b$ 来线性拟合 F 样本点 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，…，$\underset{\sim}{A}_n$ 的置信度越高。由于在 $\underset{\sim}{Q} \in \mathscr{F}(R^2)$和 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，…，$\underset{\sim}{A}_n \in \mathscr{F}(R^2)$给定的前提下，$F$ 相关系数 R 的值取决于直线 l：$y=ax+b$ 的系数 a 和 b。因此可视 R 为 a，b 的函数。记为

$$R = R(a, b).$$

可以证明，在 $\mu_{\underset{\sim}{A}_1}$，$\mu_{\underset{\sim}{A}_2}$，…，$\mu_{\underset{\sim}{A}_n}$连续时，$R$ 是 a，b 的连续函数（证明见原文定理 1）。

定义 6.3.5 设 R 是相关系数，则 F 线性拟合相关系数定义为

$$R^* = R(a^*, b^*) = \sup_{a, b} R(a, b).$$

此时，直线 $l^*: y=a^*x+b$ 定义为 F 误差标准 $\underset{\sim}{Q}$ 的要求下，F 样本点 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，…，$\underset{\sim}{A}_n$ 的 F 线性拟合直线。

至于 R^*，a^*，b^* 的计算可以使用数值计算的方法。例如，从实际问题出发，可以大致确定 a^* 和 b^* 的区间范围 $[\underline{a}, \overline{a}]$ 和 $[\underline{b}, \overline{b}]$，然后将 $[\underline{a}, \overline{a}]$ 和 $[\underline{b}, \overline{b}]$ 分别等分为 M 和 N 份，计算每一个节点 (a_i, b_j) 的 F 相关系数 R_{ij}，$(i=0, 1, \cdots, M; j=0, 1, \cdots, N)$。比较 R_{ij}，使 $R_{i_0 j_0}$ 取得最大值，则对应有近似解：

$$a^* \approx a_{i_0}, \quad b^* \approx b_{j_0}, \quad R^* \approx R(a_{i_0}, b_{j_0}).$$

在测得 n 个 F 样本点 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，…，$\underset{\sim}{A}_n$ 以后，R^* 的大小由 F 误差标准集 $\underset{\sim}{Q}$ 所确定，误差标准越高，即误差范围要求越小，则 R^* 的值也越小；反之，则 R^* 的值越大。另一方面，在 F 误差标准集 $\underset{\sim}{Q}$ 给定后，$R^*=R(a^*, b^*)$ 值的大小，反映了 F 线性拟合直线 $l^*: y=a^*x+b^*$ 对 F 样本点进行拟合的满意程度。R^* 的值越大，F 线性拟合的可信度越高，R^* 的值越小，则 F 线性拟合的可信度越低。因此，R^* 可视为 F 样本点和 F 线性拟合直线之间的相关性指标，它反映了线性拟合的满意程度，即可信度.

有了 R^* 及 l^*，我们可利用它们来进行预测：对任给的 $\hat{x} \in X$，值 $\hat{y}=a^*\hat{x}+b^*$ 称为估计值。由于 F 性，$\hat{y}$ 的取值应有一个误差范围，这个误差范围为 $\pm \mu_{\underset{\sim}{Q}}^{-1}(R^*)$，则 y 的取值范围为

$$(\hat{y}-\mu_{\underset{\sim}{Q}}^{-1}(R^*), \ \hat{y}+\mu_{\underset{\sim}{Q}}^{-1}(R^*)).$$

其中 $\mu_{\underset{\sim}{Q}}^{-1}$ 为 $\underset{\sim}{Q}$ 的隶属函数的反函数，并规定

$$\underset{\sim}{\mu_0}^{-1}(1)=\lim_{R\to 1-0}\underset{\sim}{\mu_0}^{-1}(R).$$

由上面讨论可知，在测得 F 样本点 $\underset{\sim}{A}_1$，$\underset{\sim}{A}_2$，…，$\underset{\sim}{A}_n$ 以后，对 F 线性拟合相关系数 $R^*=(a^*, b^*)$的求解，得到一条清晰的直线 $l^*: y=a^*x+b^*$ 去拟合 F 样本点，并且这种拟合的可信度即是 R^*。在具体计算过程中 $\underset{\sim}{Q}$ 的选取不一定用线性隶属函数，$R^*=R(a^*, b^*)$的计算也可以用其它更有效的方法，而且整个过程可以在计算机上完成。不难看出，此法可拓广至 R^n 中 F 样本点及非线性拟合方面。

第七章

F聚类分析

我们所研究所接触到客观世界的对象(我们称之为样本)是十分复杂的，但是人们总是希望对它们加以条理化，从而产生一个基本的科学问题：如何将这些对象加以分类。换句话说，对于给定的 n 个对象及其属性的观测，根据内在的特征对它们进行分类。这是一类相当广泛的课题。在此，我们仅讨论分类问题的一个侧面：不分明聚类分析。

本章将在回顾经典的分类问题的基础上介绍不分明聚类分析，并给出一定量的例题。学会这种方法，将有助于解决实际问题。

§1 经典的聚类分析

在数理统计中把按一定要求和规律对事物进行分类的方法叫做聚类分析。这种方法由于应用十分广泛，近十多年来发展很快。

例如，在商业部门中，往往需要将商店按其经销的商品分成若干类型，如食品店、五金店和服装店等等。又如，在生物学中需要将动物分为两栖类、爬虫类等等。这种诸如此类的将样本按它们的某些特性进行分类的过程，我们称为聚类分析。

设有 n 个待分类的样本：x_1，x_2，…，x_n。记其集为

$$X=\{x_1,x_2,\cdots,x_n\}。$$

每个样本均具有S种特性，其数量化指标为$y_1, y_2, \cdots, y_S$，其中y_i表示描述样本第i个特性的数，我们称这组数为样本的S个指标。用x_{ij}表示第i个样本的第j个指标。从而得到各种指标结果列于表1。

表 1

样本 \ 指标	y_1	y_2	y_3	$\cdots$	y_S
x_1	x_{11}	x_{12}	x_{13}	$\cdots$	x_{1S}
x_2	x_{21}	x_{22}	x_{23}	$\cdots$	x_{2S}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	x_{n1}	x_{n2}	x_{n3}	$\cdots$	x_{nS}

从表1中可见样本x_i可用行矩阵$(x_{i1}, x_{i2}, \cdots, x_{iS})$来表示：$x_i = (x_{i1}, x_{i2}, \cdots, x_{iS})$。

现在的问题是，根据上述的观测结果，对样本集X进行一个划分或对x_1，x_2，$\cdots$，x_n进行一个分类。即寻求一种方案，将X进行一次划分。所谓划分，是指X上的非空子集类$G = \{G_1, G_2, \cdots, G_m\}$，$m \leqslant n$，满足

$$X = \bigcup_{i=1}^{m} G_i, \quad G_i \cap G_j = \phi, \quad (i \neq j).$$

并称G_i为X的一个类。

可见，一个聚类过程，就是寻求一种方法以求出G_1，G_2，$\cdots$，G_m。要找到该方法，就要有一个标准，用这个标准来衡量样本之间的接近或相似程度，以区别样本，便于分类。最简单的接近程度尺度是广义距离。

样本

$$x_i = (x_{i1}, x_{i2}, \cdots, x_{iS})$$

和

$$x_j=(x_{j1}, x_{j2}, \cdots, x_{jS})$$

的距离用 $d_{ij}=\|x_i-x_j\|$ 来表示，并规定

$$d_{ij}^2=\sum_{K=1}^{S}(x_{iK}-x_{jK})^2.$$

这个距离综合的反映了两样本之间的每个指标间差异的大小，一般称其为欧氏距离。

例 7.1.1 设抽了七个样品，每个只测了一个指标，它们是 1，2，5，7，9，10，11。试将它们聚类。

解 为了将这七个样品分类，自然考虑它们之间的相互距离，依距离大小来分类。距离小于某个给定值的归为一类，而距离大的分于不同类中。样品之间的距离列于表2中。

这里是一般的特例：$n=7$，$S=1$。

表 2

d_{ij}	1	2	5	7	9	10	11
1	0						
2	1	0					
5	4	3	0				
7	6	5	2	0			
9	8	7	4	2	0		
10	9	8	5	3	1	0	
11	10	9	6	4	2	1	0

由表可见，最短距离为 1。从而应将两样品间距离为 1 的样品归为 1 类。于是将 1 与 2 归为一类，9，10，11 分为 1 类，得四类：

$$\{1, 2\}, \{5\}, \{7\}, \{9, 10, 11\}.$$

问题并未完全解决，因为这里是按距离为 1 来区分的，当它取其它值时分类仍可进行。但要再归类，就要涉及类与类之间的距离概念。如类{1, 2}与{9, 10, 11}之间的距离是什么？为此我们把类看成新的样本，用以代替该类中原有的各样本，并以后者的各项指标的平均值作为新样本的该项指标。如类{9, 10, 11}看成新样本，其指标为{10}，同样{1, 2}为{1.5}。于是类{1, 2}与类{9, 10, 11}之间的距离为 8.5 。如此对于类{1, 2}, {5}, {7}, {9, 10 , 11} 可算出其间距离，列于表 3 中。

表 3

	{1, 2}	{5}	{7}	{9, 10, 11}
{1, 2}	0			
{5}	3.5	0		
{7}	5.5	2	0	
{9, 10, 11}	8.5	5	3	0

由表可见，最近距离为 2。从而可将 5 与 7 归为一类：{5, 7}。这时 X 划分为三类，即{1,2}, {5,7}, {9,10,11}.

重复上述过程，类{5,7}与类{1, 2}及{9, 10, 11}之间距离分别为 4.5 及 4。若以最近距离 4 为准，则 X 划分为两类：

{1, 2}, {5, 7, 9, 10, 11}.

若提高距离标准，则最后 X 成为一类。聚类过程结束。

从这个例子可以看出聚类分析的一个大体轮廓。但要实现对样本的分类，就必须建立一个普遍的法则。在传统的聚类分析中常应用两大类方法：系统聚类法与逐步聚类法。下面我们扼要的介绍这两种方法。

7.1.1 系统聚类法

该方法的步骤如下：

(1) 计算各样本间的距离，将距离最近的两样本归为一类。

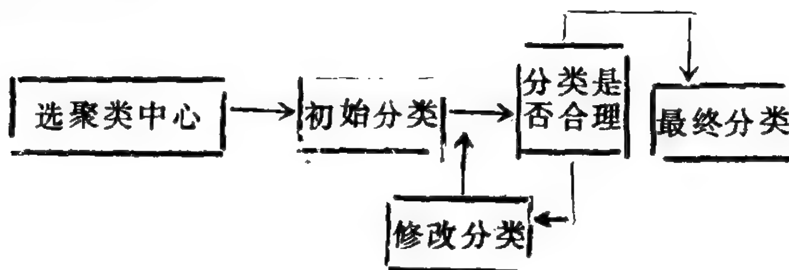
(2) 把拼成的类看成新的样本，用以代替该类中的所有样本，并以后者各项指标的平均值，作为新样本的各项指标。

(3) 反复进行步骤 1°与 2°，直至所有样本都归为一类.

这种方法的具体实施已在例 7.1.1 中指明。系统分类法的优点是一次形成分类，缺点是计算量较大。

7.1.2 逐步聚类法

用系统聚类法聚类，样本一旦划到某个类以后就不变了，这要求分类的方法比较准确．此外，由于它的计算量较大，往往使人难以胜任。因此，是否能先给一个粗糙的初始分类，然后用某种原则进行修改，直至分类比较合理为止，采用这种思想产生的聚类法叫做逐步聚类法或动态聚类法．为了得到初始分类，有时设法选择一些凝聚点(聚类中心)，让样本按某种原则向凝聚点聚类．逐步分类法大体可用如下框图来表示：



框图的每一部分均有多种方法，这些方法按框图组合就会得到各种动态聚类法(详见[2])。

一般说来其步骤如下：

(1) 选择一批聚类中心。所谓聚类中心就是某一样本的核心，它可能并不是任何一个样本，但它的指标却反映了该类的特征，因而可以看成是某一类的标准假想样本。一般，它所对应的各个指标是该类样本所对应的指标的平均值。

(2) 我们将样本向最近的聚类中心聚类，从而将样本分类。

(3) 根据分类结果找出各类的新聚类中心，它的各项指标即为该类所有样本的相应指标的平均值。然后计算这前后两组聚类中心的差异，如果差异大于某个阈值，即认为分类不合理。

(4) 修改分类，即以新的聚类中心代替旧的，我们反复进行分类。判断合理性和修改聚类中心，直至前一次聚类中心与后一次的聚类中心的差异小于某个阈值，即认为分类合理，从而分类过程结束.最后一次得到的分类就是最终分类.

例 7.1.2 从 21 个工厂抽了同类产品，每个产品测了两个指标。欲将各厂的质量情况进行分类。测得的数据如下(已作了适当变换)：

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
x_1	0	0	2	2	4	4	5	6	6	7	−4	−2	−3	−3	−5	1	0	0	−1	−1	−3
x_2	6	5	5	3	4	3	1	2	1	0	3	2	2	0	2	1	−1	−2	−1	−3	−5

这里样本容量 $n=21$，而 $S=2$。我们利用系统聚类法中最短距离法分类，则分成以下诸类：{1，2，3，4，5，6}，{7，8，9，10}，{11，12，13，14，15}，{17,18,19,20}，{16}，{21}。

其次用逐步聚类法求解。

(1) 选取聚类中心。我们按某种方法或人为的选取。这里取 x_{18}, x_{17}, x_6, 三个聚类中心。

(2) 初始分类. 样本之间采用欧氏距离 $d_{ij}=\|x_i-x_j\|$。样本按最小距离归类。不难算得样本分为三类(类内数字 i 表示样本 x_i):

$$G_1^0=\{3, 4, 5, 6, 7, 8, 9, 10\},$$

$$G_2^0=\{2, 11, 12, 13, 14, 15\},$$

$$G_3^0=\{16, 17, 18, 19, 20, 21\},$$

由于 $d_{1,6}=d_{1,13}=\sqrt{25}$, 故 x_1 暂不归类。

(3) 修改分类。以各类重心作为新的聚类中心。重心定义为样本均值。因此三个类重心为

$$G_1^0: (4.5, 2.375),$$

$$G_2^0: (-2.83, 2.33),$$

$$G_3^0: (-0.67, -1.83).$$

它们不同于原来聚类中心 $x_6=(4, 3)$, $x_{13}=(-3, 2)$, $x_{17}=(0, -1)$。故将重心看成新的聚类中心, 将样本重新归类, 分类结果为

$$G_1^1=\{3, 4, 5, 6, 7, 8, 9, 10\},$$

$$G_2^1=\{1, 2, 11, 12, 13, 14, 15\},$$

$$G_3^1=\{16, 17, 18, 19, 20, 21\}.$$

重新计算三个类的重心:

$$G_1^1: (4.5, 2.375),$$

$$G_2^1: (-2.43, 2.86),$$

$$G_3^1: (-0.67, -1.83).$$

其中 G_2^0 与 G_2^1 的重心不同, 故以 G_1^1, G_2^1, G_3^1 的重心作为

新的聚类中心，重新分类。分类结果同 G_1^1，G_2^1，G_3^1 。因此 G_1^1，G_2^1，G_3^1是最终分类。

从以上步骤和例子看出，在每一步骤中所使用的方法是多种多样的，因而聚类结果也可能不同。另外，由此也产生一个问题，是否有一个最佳的方案呢？回答是肯定的。但我们的目的是探讨不分明聚类分析，因而不再深究。

§2 利用不分明关系的聚类法

众所周知，当两组事物之间的关系不宜用"有"或"无"作肯定或否定的回答时，可以考虑用不分明关系来加以描述。其中的不分明等价关系却可以用来作为不分明聚类的依据。

7.2.1 不分明关系

集合 X 到集合 Y 中的一个不分明关系 $\underset{\sim}{R}$，是直积空间 $X\times Y$ 中的一个不分明子集合。集合 X 到集合 X 中的不分明关系，称为集合 X 上的不分明关系。

一般说来，只要给出直积空间 $X\times Y$ 中的不分明集合 $\underset{\sim}{R}$ 的隶属函数 $\mu_{\underset{\sim}{R}}(x, y)$，集合 X 到集合 Y 的不分明关系 $\underset{\sim}{R}$ 就确定了。

正如通常的关系可用矩阵来表示一样，不分明关系可用不分明矩阵来表示。注意，这里涉及的集合都是有限集。$\mu_{\underset{\sim}{R}}(x, y)$表示集合 X 中的 x 和集合 Y 中的 y 从属于 F 关系 $\underset{\sim}{R}$ 的程度。

例 7.2.1 某家庭中子女与父母外貌的相似关系 $\underset{\sim}{R}$ 为一不分明关系，可表示为

$\underset{\sim}{R}$	父	母
子	0.9	0.2
女	0.1	0.7

亦可用矩阵来表示，即

$$\underset{\sim}{R}=\begin{pmatrix}0.9 & 0.2\\ 0.1 & 0.7\end{pmatrix}.$$

其中 $\mu_{\underset{\sim}{R}}$(父，子)$=0.9$，$\mu_{\underset{\sim}{R}}$(父，女)$=0.1$，

$\mu_{\underset{\sim}{R}}$(母，子)$=0.2$，$\mu_{\underset{\sim}{R}}$(母，女)$=0.7$。

一般，若 $X=\{x_1, x_2, \cdots, x_n\}$ 和 $Y=\{y_1, y_2, \cdots, y_S\}$均为有限集。$X\times Y$ 中的 F 关系可表为如下的 F 矩阵：

$$\begin{pmatrix}\mu_{\underset{\sim}{R}}(x_1, y_1), & \mu_{\underset{\sim}{R}}(x_1, y_2), & \cdots, & \mu_{\underset{\sim}{R}}(x_1, y_S)\\ \mu_{\underset{\sim}{R}}(x_2, y_1), & \mu_{\underset{\sim}{R}}(x_2, y_2), & \cdots, & \mu_{\underset{\sim}{R}}(x_2, y_S)\\ \cdots & \cdots & & \cdots\\ \mu_{\underset{\sim}{R}}(x_n, y_1), & \mu_{\underset{\sim}{R}}(x_n, y_2), & \cdots, & \mu_{\underset{\sim}{R}}(x_n, y_S)\end{pmatrix}$$

或简记为 $\underset{\sim}{R}=[\mu_{ij}]$，即

$$\underset{\sim}{R}=\begin{pmatrix}\mu_{11} & \mu_{12} & \cdots & \mu_{1S}\\ \mu_{21} & \mu_{22} & \cdots & \mu_{2S}\\ \cdots & \cdots & \cdots & \cdots\\ \mu_{n1} & \mu_{n2} & \cdots & \mu_{nS}\end{pmatrix}$$

$$0\leqslant\mu_{ij}\leqslant1,\quad i=1, 2, \cdots, n,$$
$$j=1, 2, \cdots, S.$$

它是一个 n 行 S 列矩阵。

F 矩阵虽然在形式上和矩阵(元素值在0与1之间)一致，但其含义不同，且其间运算关系不同。下面复习一下 F

矩阵间的运算关系及复合。

定义 7.2.1 设 $\underset{\sim}{A}=[a_{ij}]$，$\underset{\sim}{B}=[b_{ij}]$为两个 F 矩阵，则称

$$[c_{ij}]=[\max(a_{ij},\ b_{ij})]=[a_{ij}\vee b_{ij}],$$
$$[c_{ij}]=[\min(a_{ij},\ b_{ij})]=[a_{ij}\wedge b_{ij}],$$
$$[c_{ij}]=[1-a_{ij}],$$
$$[c_{ij}]=[\max_k\min(a_{ik},\ b_{kj})]=\left[\bigvee_k(a_{ik}\wedge b_{kj})\right]$$

分别为 A 与 B 的并、交、A 的逆及复合，并且相应记为 $\underset{\sim}{C}=\underset{\sim}{A}\cup\underset{\sim}{B}$，$\underset{\sim}{C}=\underset{\sim}{A}\cap\underset{\sim}{B}$，$\underset{\sim}{C}=\overline{\underset{\sim}{A}}$ 及$\underset{\sim}{C}=\underset{\sim}{A}\circ\underset{\sim}{B}$。

例 7.2.2 若有

$$\underset{\sim}{A}=\begin{bmatrix}0.8 & 0.5\\0.3 & 0.7\end{bmatrix},\qquad \underset{\sim}{B}=\begin{bmatrix}0.5 & 0.3\\0.4 & 0.8\end{bmatrix}$$

则

$$\underset{\sim}{A}\cup\underset{\sim}{B}=\begin{bmatrix}0.8 & 0.5\\0.3 & 0.7\end{bmatrix}\cup\begin{bmatrix}0.5 & 0.3\\0.4 & 0.8\end{bmatrix}$$
$$=\begin{bmatrix}0.8\vee0.5 & 0.5\vee0.3\\0.3\vee0.4 & 0.7\vee0.8\end{bmatrix}=\begin{bmatrix}0.8 & 0.5\\0.4 & 0.8\end{bmatrix},$$

$$\underset{\sim}{A}\cap\underset{\sim}{B}=\begin{bmatrix}0.8\wedge0.5 & 0.5\wedge0.3\\0.3\wedge0.4 & 0.7\wedge0.8\end{bmatrix}=\begin{bmatrix}0.5 & 0.3\\0.3 & 0.7\end{bmatrix},$$

$$\overline{\underset{\sim}{A}}=\begin{bmatrix}1-0.8 & 1-0.5\\1-0.3 & 1-0.7\end{bmatrix}=\begin{bmatrix}0.2 & 0.5\\0.7 & 0.3\end{bmatrix},$$

$$\underset{\sim}{A}\circ\underset{\sim}{B}=\begin{bmatrix}0.8 & 0.5\\0.3 & 0.7\end{bmatrix}\circ\begin{bmatrix}0.5 & 0.3\\0.4 & 0.8\end{bmatrix}$$
$$=\begin{bmatrix}(0.8\wedge0.5)\vee(0.5\wedge0.4), & (0.8\wedge0.3)\vee(0.5\wedge0.8)\\(0.3\wedge0.5)\vee(0.7\wedge0.4), & (0.3\wedge0.3)\vee(0.7\wedge0.8)\end{bmatrix}$$

$$= \begin{pmatrix} 0.5 & 0.5 \\ 0.4 & 0.7 \end{pmatrix}.$$

一般　$\underset{\sim}{A} \circ \underset{\sim}{B} \neq \underset{\sim}{B} \circ \underset{\sim}{A}$。

定义 7.2.2　设 $\underset{\sim}{R}$ 是 X 上的一个 F 关系：$\underset{\sim}{R} = [\mu_{ij}]$。如果它满足

(1) 自返性：$\mu_{ii} = 1$，$i = 1, 2, \cdots, n$。

(2) 对称性：$\mu_{ij} = \mu_{ji}$，$i \neq j$，$i, j = 1, 2, \cdots, n$。

则称 $\underset{\sim}{R}$ 为不分明相容关系。若 $\underset{\sim}{R}$ 再满足

(3) 传递性：$\underset{\sim}{R} \circ \underset{\sim}{R} \subseteq \underset{\sim}{R}$，

则称 $\underset{\sim}{R}$ 为 F 等价关系。

在定义中，条件(1)说明 $\underset{\sim}{R}$ 阵中对角线上元素皆为 1。即每个元素 $x \in X$ 与自身从属于 F 关系 $\underset{\sim}{R}$ 的程度为 1。例如，相象关系就具有自返性，而仇敌关系就不具有自返性；条件(2)指明 x_i 与 $x_j \in X$ 从属于 $\underset{\sim}{R}$ 的程度和 x_j 与 x_i 隶属于 $\underset{\sim}{R}$ 的程度一样。例如，相象关系就具有对称性，而相爱关系就不具有对称性；条件(3)是说，对任意的 X 中样本 x_i，x_j，x_k，有

$$\mu_{\underset{\sim}{R}}(x_i, x_k) \geqslant \min[\mu_{\underset{\sim}{R}}(x_i, x_j), \mu_{\underset{\sim}{R}}(x_j, x_k)]$$

即 x_i 与 x_k 隶属于 $\underset{\sim}{R}$ 的程度不小于 x_i 与 x_j 隶属于 $\underset{\sim}{R}$ 的程度与 x_j 和 x_k 隶属于 $\underset{\sim}{R}$ 的程度中较小的那一个。例如，“小得多”关系具有传递性而相象关系就不具有传递性。

例 7.2.3　对五个工厂同类产品的四个指标进行测量，得到如下结果：

Ⅰ：(5, 5, 3, 2)；Ⅱ：(2, 3, 4, 5)；

Ⅲ：(5, 5, 2, 3)；Ⅳ：(1, 5, 3, 1)；

V：(2, 4, 5, 1)。

并以某种方式得到其间的相似关系 $\underset{\sim}{R}$ 阵：

$$\underset{\sim}{R}=\begin{pmatrix} 1 & 0.81 & 0.98 & 0.89 & 0.87 \\ 0.81 & 1 & 0.79 & 0.79 & 0.21 \\ 0.98 & 0.79 & 1 & 0.83 & 0.79 \\ 0.89 & 0.79 & 0.83 & 1 & 0.93 \\ 0.87 & 0.21 & 0.79 & 0.93 & 1 \end{pmatrix}$$

由于 $\underset{\sim}{R}$ 中对角线上元素全为 1，所以具有自返性；其次显然 $\underset{\sim}{R}$ 也具有对称性。所以 $\underset{\sim}{R}$ 是一个相容关系。但是 $\underset{\sim}{R}$ 却不具有传递性，事实上，

$$\underset{\sim}{R}\circ\underset{\sim}{R}=\begin{pmatrix} 1 & 0.81 & 0.98 & 0.89 & 0.89 \\ 0.81 & 1 & 0.81 & 0.81 & 0.81 \\ 0.98 & 0.81 & 1 & 0.89 & 0.87 \\ 0.89 & 0.81 & 0.89 & 1 & 0.93 \\ 0.89 & 0.81 & 0.87 & 0.93 & 1 \end{pmatrix}$$

显然，$\underset{\sim}{R}\circ\underset{\sim}{R}\not\subseteq\underset{\sim}{R}$。因此 $\underset{\sim}{R}$ 不是一个不分明等价关系。

是否一个不分明相容关系可以改造为一个 F 等价关系呢？下面的定理回答了这一问题。

定义 7.2.3 若 $\underset{\sim}{R}$ 是集合 X 上的不分明关系，我们称

$$\underset{\sim}{R}^n=\underbrace{\underset{\sim}{R}\circ\underset{\sim}{R}\circ\cdots\circ\underset{\sim}{R}}_{n\text{个}}$$

为集合 X 上 $\underset{\sim}{R}$ 的 n 级 F 关系或 $\underset{\sim}{R}$ 的 n 次幂。其中 n 为任一正整数。

定理 7.2.1 设 $\underset{\sim}{R}$ 是集合 X 上的一个 F 相容关系，则

$$\lim_{n\to\infty}\underset{\sim}{R}^n=\underset{\sim}{R}^\infty$$

必存在，且是一个 F 等价关系。

证 设 $\underset{\sim}{R}=(\mu_{ij})$，因 $\mu_{ii}=1$，故 $\mu_{ii}\wedge\mu_{ij}=\mu_{ij}$，于是 $\mu_{ij}\leqslant\bigvee_{k=1}^{n}(\mu_{ik}\wedge\mu_{kj})=(\underset{\sim}{R}\circ\underset{\sim}{R})_{ij}$，对一切 i，j成立。从而 $\underset{\sim}{R}^2=\underset{\sim}{R}\circ\underset{\sim}{R}\geqslant\underset{\sim}{R}$。

同理可知，$\underset{\sim}{R}\leqslant\underset{\sim}{R}^2\leqslant\underset{\sim}{R}^3\leqslant\cdots\leqslant\underset{\sim}{R}^n\leqslant\cdots\leqslant E$。

由此得 $\underset{\sim}{R}^\infty=\lim_{n\to\infty}\underset{\sim}{R}^n=\bigcup_{n=1}^{\infty}\underset{\sim}{R}^n$

又因 $\underset{\sim}{R}\circ\underset{\sim}{R}^\infty=\underset{\sim}{R}\circ\left(\bigcup_{n=1}^{\infty}\underset{\sim}{R}^n\right)=\bigcup_{n=1}^{\infty}(\underset{\sim}{R}\circ\underset{\sim}{R}^n)$

$$=\bigcup_{n=1}^{\infty}\underset{\sim}{R}^{n+1}=\underset{\sim}{R}^\infty$$

同理 $\underset{\sim}{R}^m\circ\underset{\sim}{R}^\infty=\underset{\sim}{R}^\infty$。于是

$$\underset{\sim}{R}^\infty\circ\underset{\sim}{R}^\infty=\left(\bigcup_{n=1}^{\infty}\underset{\sim}{R}^n\right)\circ\underset{\sim}{R}^\infty=\bigcup_{n=1}^{\infty}(\underset{\sim}{R}^n\circ\underset{\sim}{R}^\infty)$$

$$=\bigcup_{n=1}^{\infty}\underset{\sim}{R}^\infty=\underset{\sim}{R}^\infty.$$

从而得到 $\underset{\sim}{R}^\infty$具有传递性，加上所设条件可知 $\underset{\sim}{R}^\infty$ 为一 F 等价关系。

定理 7.2.2 若 $\underset{\sim}{R}$ 是 $n\times n$ 阶反身、对称不分明矩阵，则存在 k，使

$$\underset{\sim}{R}^k=\underset{\sim}{R}^\infty.$$

证 $\underset{\sim}{R}$ 由至多 n^2个数组成，$\underset{\sim}{R}$ 的任何次幂都由其中一部分数组成.由n^2个数组成的 $n\times n$ 阶矩阵至多有$(n^2)^{n^2}$个。故当 $k>(n^2)^{n^2}$时，在$\{\underset{\sim}{R}^m\}(m\geqslant k)$中至少有两个相同．又因 $\underset{\sim}{R}\leqslant\underset{\sim}{R}^2\leqslant\cdots\leqslant\underset{\sim}{R}^m\leqslant\cdots\leqslant E$，故若 $\underset{\sim}{R}^i=\underset{\sim}{R}^j$，$(i<j)$，则对

任意的 $l \geqslant i$ 皆有 $\underset{\sim}{R}^l = \underset{\sim}{R}^i$。

1974 年 Dunn证明了[10]，存在 $k \leqslant n$ 使 $\underset{\sim}{R}^k = \underset{\sim}{R}^\infty$成立。据此，取 $\underset{\sim}{R}$ 的二次幂的方法：若在

$$\underset{\sim}{R}^2,\ \underset{\sim}{R}^4,\ \underset{\sim}{R}^8,\ \cdots$$

中某一步有

$$\underset{\sim}{R}^k = \underset{\sim}{R}^{2k} = \underset{\sim}{R}^*$$

则 $\underset{\sim}{R}^*$便是一个 F 等价关系。如此一来，就将 $\underset{\sim}{R}$ 改造为一个 F 等价矩阵。一般所需要的步数$\leqslant \lg n/\lg 2$。

例 7.2.4 继续例 7.2.3，改造 $\underset{\sim}{R}$ 为一个 F 等价关系。

因为

$$\underset{\sim}{R}^2 \circ \underset{\sim}{R}^2 = \underset{\sim}{R}^4 = \begin{pmatrix} 1 & 0.81 & 0.98 & 0.89 & 0.89 \\ 0.81 & 1 & 0.81 & 0.81 & 0.81 \\ 0.98 & 0.81 & 1 & 0.89 & 0.89 \\ 0.89 & 0.81 & 0.89 & 1 & 0.93 \\ 0.89 & 0.81 & 0.89 & 0.93 & 1 \end{pmatrix}$$

且 $\underset{\sim}{R}^8 = \underset{\sim}{R}^4 \circ \underset{\sim}{R}^4 = \underset{\sim}{R}^4$。故 $\underset{\sim}{R}^4 = \underset{\sim}{R}^\infty$是一个 F 等价关系。

7.2.2 利用 F 等价关系进行分类

若在样本集 $X = \{x_1, x_2, \cdots, x_n\}$中样本间已建立起一个$F$等价关系 $\underset{\sim}{R}$；$\underset{\sim}{R} = [\mu_{ij}]$，那么就可以根据 μ_{ij} 的大小分类。

定义 7.2.4 对 X 中任意两个样本 x_i, x_j，若

$$\mu_{ij} = \mu_{\underset{\sim}{R}}(x_i, x_j) \geqslant \lambda$$

其中 $\lambda \in [0,1]$为一定数，则 x_i 与 x_j 属于同一类，否则不属于同一类。

阈值 λ 的选取将根据实际需要而定，归类过程是唯一的。

例 7.2.5 试将例7.2.3的$X = \{ \mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV}, \mathrm{V} \}$进行分

类。

解 因为$\underset{\sim}{R}^4$是一个F等价关系，故利用$\underset{\sim}{R}^4$进行分类。取$\lambda=1$时。X被分成五类：

$$\{Ⅰ\},\{Ⅱ\},\{Ⅲ\},\{Ⅳ\},\{Ⅴ\}.$$

即每个样本自成一类。

若取$\lambda=0.90$，则X被分成三类：

$$\{Ⅰ,Ⅲ\},\{Ⅱ\},\{Ⅳ,Ⅴ\}.$$

若取$\lambda=0.85$，则X被分为两类：

$$\{Ⅰ,Ⅲ,Ⅳ,Ⅴ\},\{Ⅱ\}$$

若取$\lambda=0.80$，则X被分为一类，即

$$\{Ⅰ,Ⅱ,Ⅲ,Ⅳ,Ⅴ\}.$$

由例可见，λ取值不同可得不同的结论。从而在分类过程中恰当调整λ值以得到合适的分类。

整个过程的大体步骤可概括如下：

（1）按照待聚类样本的特征建立一个样本间的F相容关系。对多指标的样本往往先将各指标数量化，并将它们进行适当的规格化处理，然后用某种方法来建立各样本间的F关系。这样得到的关系是F相容关系$\underset{\sim}{R}$。

（2）通过合成(复合)运算，求得对应的F等价关系$\underset{\sim}{R}^*$.

（3）取不同的λ值得到不同的分类结果，再根据实际需要选择合理的分类。

如果我们引进R_λ矩阵，则分类将变得更为清晰。

定义 7.2.2 对任意的$\lambda\in[0,1]$,矩阵$R_\lambda=[\lambda_{ij}]$,其中

$$\lambda_{ij}=\begin{cases}1, & 当\ \mu_{ij}\geqslant\lambda\\ 0, & 当\ \mu_{ij}<\lambda\end{cases}$$

称R_λ为$\underset{\sim}{R}=[\mu_{ij}]$的$\lambda$截矩阵。

例 7.2.6 若

$$\underset{\sim}{R}=\begin{pmatrix}1 & 0.4 & 0.6 & 0.9\\ 0.4 & 1 & 0.7 & 0.5\\ 0.6 & 0.7 & 1 & 0.3\\ 0.9 & 0.5 & 0.3 & 1\end{pmatrix}$$

则取 $\lambda=0.6$，0.3 时的 λ 截矩阵 R_λ 为：

$$R_{0.6}=\begin{pmatrix}1 & 0 & 1 & 1\\ 0 & 1 & 1 & 0\\ 1 & 1 & 1 & 0\\ 1 & 0 & 0 & 1\end{pmatrix}.$$

$$R_{0.3}=\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{pmatrix}.$$

在 R_λ 中，我们约定若第 i 列与第 j 列完全一样，则认为样本 x_i 与 x_j 归为一类。如此，当 λ 值由 1 下降至 0 时，整个分类过程结束。

例 7.2.7 设 $X=\{x_1,x_2,x_3,x_4\}$上的等价关系 $\underset{\sim}{R}$ 为

$$\underset{\sim}{R}=\begin{pmatrix}1 & 0.4 & 0.9 & 0.6\\ 0.4 & 1 & 0.4 & 0.4\\ 0.9 & 0.4 & 1 & 0.6\\ 0.6 & 0.4 & 0.6 & 1\end{pmatrix}$$

当 λ 取值为 1,0.8,0,6,0.4 时，有

$$R_1=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$ X 被分为四类：$\{x_1\},\{x_2\},\{x_3\},\{x_4\}$.

$$R_{0.8}=\begin{pmatrix}1&0&1&0\\0&1&0&0\\1&0&1&0\\0&0&0&1\end{pmatrix}.$$ X 被分为三类：$\{x_1,x_3\},\{x_2\},\{x_4\}$.

$$R_{0.6}=\begin{pmatrix}1&0&1&1\\0&1&0&0\\1&0&1&1\\1&0&1&1\end{pmatrix}.$$ X 被分为两类：$\{x_1,x_3,x_4\},\{x_2\}$.

$$R_{0.4}=\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}.$$ X 为一类：$\{x_1,x_2,x_3,x_4\}$ 聚类过程结束.

7.2.3 关于建立相似关系的一点说明

聚类分析在多元统计分析中所处的地位是相当重要的，由于它能够解决许多实际问题而受到重视．和多元分析中其它方法相比，它显得粗糙，理论上很不完善．但正是这一点，说明它有着广阔的前景．

聚类分析的中心思想是要找出相似关系，依此相似关系为准则将变量或样品分类．目前常用的相似关系，在样品之间有相似系数和距离的各种定义，如明考斯基距离、切比雪夫距离、马氏距离、兰氏距离以及对有序尺度和名义尺度所定义的距离等等．在变量之间分类的相似系数，有夹角余弦、

相关系数、指数相似系数、非参数相似系数、列联系统、连关系数等等(详见[31])。有了相似关系，聚类就不困难了。

对于 F 聚类分析，也有类似的情形. 目前已有不少文章出现，使用各种相似系数构成不分明矩阵进行聚类。常见的有数量积、相关系数、最大最小方法、算术平均最小方法、几何平均最小方法、绝对值倒数法及主观评定法等。

如在文[44]中利用数量积法就啤酒大麦育种问题进行探讨，取得良好的效益。被分类对象的全体样本集为 $x=\{x_1,\cdots,x_{10}\}$，其 x_i，$i=1,2,\cdots,10$，为一个样本，$x_i=\{x_{i1},\cdots,x_{im}\}$表示第 i 个品种在 m 个不同地点的千粒重，是指标集。而 $x_{ik}(k=1,\cdots,m)$表示第 i 个品种在第 k 个地点的千粒重。因此取得 50 个原始数据：

$$x_{ik},\quad i=1,\cdots,10;k=1,\cdots,5;\quad m=5.$$

它是进行分类的主要依据。

引进数量积以建立不分明矩阵 $\underset{\sim}{R}=(\mu_{ij})$，

$$\mu_{ij}=\begin{cases}1 & \text{当 } i=j,\\ \left(\sum_{K=1}^{m}x_{ik}x_{jk}\right)/M & \text{当 } i\neq j.\end{cases}\quad (m=5)$$

其中 $M=100000$，将原始数据代入 μ_{ij}，得不分明矩阵 $\underset{\sim}{R}$。由于 $\underset{\sim}{R}$ 不是不分明等价关系，进行改造，得 $\underset{\sim}{R}^2=\underset{\sim}{R}^4$，从而 $\underset{\sim}{R}^\infty=\underset{\sim}{R}^2$。对于给定的 $\lambda(0\leqslant\lambda\leqslant1)$即可进行分类。

对样品聚类的相似系数，不限于度量尺度. 对于特征尺度指标亦可作为相似系数。文[45]中介绍了一种对不同零件按其几何和工艺特征进行分类的方法，并提出加权聚类法。现简要介绍如下：

零件的相似性是指零件所具有的各种特征的相似。每种零件都具有多种特征，正是这些特征的组合，才构成区别于其它种零件的一个零件种。然而，许多零件的某些特征又可能相似或相同，这些相似或相同的特征，就构成了零件之间的相似性。一般零件特征包括几何和工艺特征。设有m个待分类的零件$X_1,\cdots,X_m$，每个零件均具有Q种特征。我们将这些特征数量化，于是每个零件就对应着一组描述它的各种特征的一组数：$y_1,\cdots,y_Q$。于是得到数据

$$X_t=(x_{t1},x_{t2},\cdots,x_{tQ}),$$

$$t=1,\cdots,m;\ Q\text{为一定数。}$$

在此，为了便于分析和使用计算机，用8×9二维矩阵表示零件特征，即$Q=72$，且将零件分为两类。于是，该二维矩阵中元素x_{ij}要么是1要么是0，即该零件有此特征与无此特征，从而得二进制特征描述的二维矩阵。

零件X_i与X_j之间的近似程度，使用广义夹角余弦这个量来刻划：

$$\cos\alpha=\frac{|(X_i,X_j)|}{\|X_i\|\cdot\|X_j\|},\quad 0\leqslant\alpha\leqslant90^\circ$$

其中$(X_i,X_j)=x_{i1}x_{j1}+x_{i2}x_{j2}+\cdots+x_{iQ}x_{jQ}$

$$\|X_i\|=\sqrt{\sum_{k=1}^{Q}x_{ik}^2},\qquad \|X_j\|=\sqrt{\sum_{k=1}^{Q}x_{jk}^2}$$

通过原始数据x_{ij}可算出$\cos\alpha$，以r_{ij}记之。从而得不分明矩阵$\underset{\sim}{R}$。剩下的分类办法与前面所述相同。

在分类时我们考虑到突出对划分零件族影响较大的特征，并用它作为分类的主要依据，从而引入加权法构成F矩

阵 $\underset{\sim}{R}^*$.

设 X_i^* 为零件具有某些方面特征的向量，$X_i^*=(x_{p1}^*,\cdots,x_{il}^*)$,($l<Q$)，是由 X_i 的一部分分量组成的子向量．那么两两对应子向量的加权相似系数 r_{ij}^* 定义为：

$$r_{ij}^*=\frac{|(X_i^*,X_j^*)|}{\|X_i^*\|\cdot\|X_j^*\|}$$

其中 $(x_i^*,x_j^*)=\sum_{k=1}^{l}W(k)x_{ik}^*x_{jk}^*)$

$$\|X_i^*\|=\sqrt{\sum_{k=1}^{l}W(k)x_{ik}^{*2}},$$

$W(k)(k=1,\cdots,l)$是加于第 k 个特征的权数，且

$\frac{1}{l}\sum_{k=1}^{l}W(k)=1$.

以 r_{ij}^* 为矩阵元素，就得到不分明矩阵 $\underset{\sim}{R}^*$.

通过以上的介绍可以看到聚类分析的主要步骤应是定义相似系数，构造不分明矩阵，并使之成为一个不分明等价关系．根据实际问题的要求选定阈值 λ．从而由 R_λ 即可分类．最后结合实际对分类予以分析说明．

相似系数的定义，有许多种，要根据实际问题选用或由实际情形加以定义．在这方面的文章较多，读者不难从模糊数学中找到．而关于经典聚类分析方面的相似系数等方面的汇总，可在文[37]中找到．

§3 软划分

本节我们介绍另外一种 F聚类方法：软划法分．它与常规的逐步聚类法有类似之处．为了叙述方便，我们先介绍常规

分类的情况，也称之为硬划分。

不论是硬划分还是软划分都要求事先确定好被分类的样本应该分成几类。以下均假定需要样本集 X 分为 C 类。

7.3.1 硬划分

设待分类的样本集合为

$$X=\{x_1,x_2,\cdots,x_n\},$$

每个样本 $x_i(i=1,2,\cdots,n)$ 有 s 个指标. 于是每个样本 x_i 可写成行矩阵

$$x_i=(x_{i1},x_{i2},\cdots,x_{is})。$$

现在我们要把样本集合 X 中的样本分为 C 类：$G_1,G_2,\cdots,G_c$，则对应每一种分法，可用一个 C 行 n 列，内中元素均为 0 或 1 的矩阵来表示。例如下列矩阵 U 就表示一种分类。

$$U=\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & \cdots & x_n \end{array} \\ \left(\begin{array}{ccccc} 0 & 0 & 1 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{array}\right) \end{array} \begin{array}{c} \\ G_1 \\ G_2 \\ G_3 \\ \vdots \\ G_c \end{array}$$

矩阵 U 中第一列中的第二个元素为 1，其余均为 0，就表示样本 x_1 属于第二类 G_2。同样第二列中的第三个元素为 1，其余均为 0，就表示样本 x_2 属于第三类 G_3，等等。

一般说来，若 $U=[u_{ij}]$为一 n 列 c 行的矩阵。且满足

(1) $u_{ij}=0$ 或 1，$i=1,2,\cdots,c$；$j=1,2,\cdots,n$。

(2) $u_{1j}+u_{2j}+\cdots+u_{cj}=1$，$j=1,\cdots,n$。

(3) $u_{i1}+u_{i2}+\cdots+u_{in}>0$，$i=1,2,\cdots,c$。

在约定

$$u_{ij}=\begin{cases}0 & \text{表示 } x_j \overline{\in} G_i, \\ 1 & \text{表示 } x_j \in G_i.\end{cases}$$

的情况下，U 矩阵对应一种分类。显然，给出一种分类就可以得到一个对应的矩阵 U。

条件(2)说明第 j 个样本 x_j 只能属于其中的一类，而不能同时属于两类以上。条件(3)说明每一类均非空。

例 7.3.1 $X=\{x_1,x_2,\cdots,x_6\}$，$c=3$，矩阵

$$U=\begin{pmatrix}1&0&0&1&0&0\\0&0&0&0&1&0\\0&1&1&0&0&1\end{pmatrix}$$

对应的分类是

$$G_1:\ \{x_1,x_4\},$$
$$G_2:\ \{x_5\},$$
$$G_3:\ \{x_2,x_3,x_6\}.$$

而矩阵

$$U=\begin{pmatrix}0&1&1&0&0&0\\0&0&0&1&0&1\\1&0&0&0&1&0\end{pmatrix}$$

对应的分类是

$$G_1:\ \{x_2,x_3\},$$
$$G_2:\ \{x_4,x_6\},$$
$$G_3:\ \{x_1,x_5\}.$$

n 个样本分到 c 个类中只有有限种分法。因而，它就对应有限个分类矩阵。这些矩阵的全体所构成的集合，我们称为划分空间。为了求得最佳分类，就需要在划分空间中挑出最佳的分类矩阵来。但是怎样才算最佳呢？这需要给出一个

标准。

首先，我们对每一类选取一个核心，称之为聚类中心。第 j 类 G_j 的聚类中心记为 V_j；

$$V_j = (v_{j1}, v_{j2}, \cdots, v_{jc}), \tag{7.3.1}$$

而

$$v_{jk} = \frac{\sum_{i=1}^{n} u_{ji} x_{ik}}{\sum_{i=1}^{n} u_{ji}}, \quad k = 1, 2, \cdots, s \tag{7.3.2}$$

它的实际含义是对应的各个指标为该类样本所对应的指标的平均值。

其次，考虑到若分类最佳，那么每一类中的所有样本与该类的聚类中心的距离平方的和应该最小。即

$$J(U, V) = \sum_{j=1}^{c} \sum_{i=1}^{n} u_{ij} \|x_i - V_i\|^2 \tag{7.3.3}$$

达最小值。

这就是最佳分类的分类标准。

怎样得到最佳分类呢？我们只好把所有可能的分类计算其 $J(U,V)$，选其最小者为之.但是这种办法实际上是难以实现的。目前，可用 PFS 聚类法予以实现(见第四节)。

7.3.2 软划分(F 划分)

在 F 领域内，样本的分类并不是那么经渭分明的，而是一个样本以某一隶属度从属于某一类，又以另一个隶属度从属于另一类。这样，样本就不是明确地属于或不属于某一类。从而，此时已无明确的类内、类间的概念。

例如，样本集 $X = \{x_1, x_2, x_3, x_4, x_5\}$ 要分为三类，我们

可以用矩阵

$$\mu=\begin{bmatrix}0.1 & 0.8 & 0.3 & 0.2 & 0\\ 0.7 & 0.05 & 0.5 & 0.1 & 0.2\\ 0.2 & 0.15 & 0.2 & 0.7 & 0.8\end{bmatrix}$$

来表示如下的软分类：

$\underset{\sim}{G}_1=\{0.1/x_1,\ 0.8/x_2,\ 0.3/x_3,\ 0.2/x_4\}$,

$\underset{\sim}{G}_2=\{0.7/x_1,\ 0.05/x_2,\ 0.5/x_3,\ 0.1/x_4, 0.2/x_5\}$,

$\underset{\sim}{G}_3=\{0.2/x_1, 0.15/x_2, 0.2/x_3, 0.7/x_4, 0.8/x_5\}$.

每一类是 X 上的一个 F 子集合。划分矩阵 μ 不同，所得分类不同，从而对应着的 F 子集合也不同。这种矩阵 μ 称为软划分矩阵，其对应的分类法称为软划分法。

一般来说，若要将样本集 $X=\{x_1,\cdots x_n\}$ 分为 c 类，则其软划分矩阵 μ 是一个 c 行 n 列矩阵 μ：

$$\mu=\begin{array}{c} \begin{matrix} x_1 & x_2 & \cdots & x_n \end{matrix} \\ \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1n}\\ \mu_{21} & \mu_{22} & \cdots & \mu_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ \mu_{c1} & \mu_{c2} & \cdots & \mu_{cn}\end{bmatrix} \end{array} \begin{matrix} \underset{\sim}{G}_1\\ \underset{\sim}{G}_2\\ \vdots\\ \underset{\sim}{G}_c\end{matrix}$$

它应具有如下性质：

(1) $0\leqslant\mu_{ij}\leqslant 1,\ \ i=1,2,\cdots,c;\ \ j=1,2,\cdots,n$,

(2) $\sum_{i=1}^{c}\mu_{ij}=1,\ \ j=1,2,\cdots,n$.

(3) $\sum_{j=1}^{n}\mu_{ij}>0,\ \ i=1,2,\cdots,c$.

这里 μ_{ij} 表示样本 x_j 从属于类 $\underset{\sim}{G}_i$ 的隶属程度。而 (2) 表示每一个样本属于各类的总隶属度为 1，(3)指明每个不分

明类非空。

显然 μ 划分矩阵不是唯一的，它有无穷多个。从而使我们考虑在如此众多的软划分矩阵中是否存在一个最佳的软划分矩阵，即是否存在一个最佳的分类呢？

为了解决这一问题，我们仿照硬划分的办法，选择聚类中心 V_i。衡量的标准仍然是样本与聚类中心的距离的平方和最小。但是在软划分的情况下，一个样本是按不同的隶属程度属于各类的，因此，应该同时考虑它与每一类的聚类中心的距离：

$$\sum_{j=1}^{c}\mu_{ij}\|x_i-V_j\|^2 \tag{7.3.4}$$

表示了样本 x_i 与各个聚类中心 V_i 的带权距离平方和，权就是 x_i 属于 G_j 的隶属程度。而总的带权距离平方和为

$$J(\mu,V)=\sum_{i=1}^{n}\sum_{j=1}^{c}\mu_{ij}\|x_j-V_j\|^2.$$

为了加强 x_i 属于各类的隶属度的对比度，在上式中再添上参数 m，即得

$$J_m(\mu,V)=\sum_{i=1}^{n}\sum_{j=1}^{c}\mu_{ij}^m\|x_i-V_j\|^2.$$

现在的问题是要得到最佳软划分，要求得适当的软划分矩阵与聚类中心。早在 1973 年 Bezdek 提出的此法中已证明了

$$V_i=\frac{\sum_{K=1}^{n}(\mu_{ik})^m x_k}{\sum_{K=1}^{n}(\mu_{ik})^m},\qquad i=1,2,\cdots,c \tag{7.3.5}$$

$$\mu_{ik}=\frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-2)}},\qquad i=1,2,\cdots,c;\quad j=1,2,\cdots,n; \tag{7.3.6}$$

其中 $d_{ij}=\|x_j-V_i\|=\left[\sum_{j=1}^{s}(x_{kj}-v_{ij})^2\right]^{\frac{1}{2}}$。

$V_i=(v_{i1},v_{i2},\cdots,v_{is})$，$i=1,2,\cdots,c$。

为了求出 V_i 及 μ_{ik}，使用反复迭代法，具体计算步骤如下：

（1）人为地先给出一个初始划分矩阵 μ_0，即人为的先将样本集 $X=\{x_1,\cdots,x_n\}$进行一次分类。它可以是硬划分也可以是软划分。

（2）根据 μ_0 和 V_i 的计算公式求出 V_i。

（3）由 V_i 和 μ_{ij} 的公式算出诸 μ_{ij}，得新的划分矩阵 μ_1：$\mu_1=(\mu_{ij})$。

（4）对于给定的小数 ε，若 $\max\{|\mu_{ij}-\mu_{ij}^0|\}$ 小于 ε，则计算停止。所得到的 μ 与 V_i 即为所求最佳软划分的划分矩阵和聚类中心。否则回到第(2)步，再根据已得的矩阵 μ 算出新的聚类中心，重复做第(3)及(4)步。

其中 ε 是根据问题的需要选取的一个小数（ε 可取为 10^{-2},10^{-3},10^{-4} 等)。

通过以上步骤得出最佳软划分矩阵后，我们常常还要求得相应的硬划分。这有以下两种方法可循。

方法一：直接方法。

即将 μ 中每一列的元素中最大者取为 1，其余取为 0.这实际上就是将样本划归隶属度大的那一类。

例 7.3.2

$$\mu=\begin{pmatrix}0.1 & 0.8 & 0.3 & 0.15 & 0 & 0.6\\0.7 & 0.05 & 0.5 & 0.1 & 0.3 & 0.2\\0.2 & 0.15 & 0.2 & 0.75 & 0.7 & 0.2\end{pmatrix}$$

对应的硬划分矩阵是

$$\mu^*=\begin{pmatrix}0 & 1 & 0 & 0 & 0 & 1\\1 & 0 & 1 & 0 & 0 & 0\\0 & 0 & 0 & 1 & 1 & 0\end{pmatrix}$$

从而将样本分为三类：$\{x_1,x_3\}$，$\{x_2,x_6\}$，$\{x_4,x_5\}$。

方法二：二次分类法.

将样本逐个与聚类中心相比较，看它与哪个聚类中心最接近就属于哪一类。即若

$$\|x_i-V_{j_0}\|=\min_j\{\|x_i-V_j\|\},$$

则 x_i 属于第 j_0 类。

一般说来，以上两种方法所得到的结果是基本一致的。

我们上面介绍的这种方法要事先知道分类的类数 c，如果所确定的分类数不合理，那么也就不可能有合理的分类，因而就要重新确定类数，重新计算。在这一点上它是不如利用不分明等价关系所进行的分类。但是它可以得出聚类中心，这往往是我们所需要的。因此，两种方法各有优点，应视具体问题的差异选择方法，才能获得满意的结果。另外，软划分求聚类中心的计算量较大，往往得上计算机。

例 7.3.3

我们下面用一个假想的数据结果进行 F聚类，以说明上述方法。

设样本集 $X=\{x_1, x_2, x_3, x_4, x_5, x_6\}$，每个样本测

量了两个指标，具体数据如下：

P_i \ x_i	x_1	x_2	x_3	x_4	x_5	x_6
1	1	2	5	7	9	10
2	0	2	4	8	7	9

先用 F 聚类法中的软划分法计算最佳划分矩阵及聚类中心。

1° 给出初始划分矩阵 U_0。取硬划分：

$$U_0=\begin{bmatrix}1 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 1 & 1 & 1\end{bmatrix}.$$

得到 G_1^0：$\{x_1,x_2\}$，G_2^0：$\{x_3,x_4,x_5,x_6\}$

2° 求聚类中心 V_1，V_2：

由公式(7.3.5)，得

$$\nu_{11}=1.5,\quad \nu_{12}=1,$$

$$\nu_{21}=7.75,\quad \nu_{22}=7,$$

从而 $$V_1=\begin{bmatrix}1.5\\1.0\end{bmatrix},\quad V_2=\begin{bmatrix}7.75\\7.00\end{bmatrix}$$

3° 计算新的划分矩阵：

先计算第 k 个样本与第 i 类间距离 d_{ik}，$i=1,2$；$k=1$，2，…，6。共有 12 个数据。它们是

$$d_{11}=1.118,\qquad d_{21}=9.724,$$

$$d_{12}=1.118,\qquad d_{22}=7.620,$$

$$d_{13}=4.610,\qquad d_{23}=4.070,$$

$$d_{14}=8.902,\qquad d_{24}=1.25,$$

$$d_{15}=9.605,\qquad d_{25}=1.118,$$

$$d_{16}=11.673,\qquad d_{26}=3.010.$$

再由 μ_{ik} 公式算得软划分矩阵 μ，其中

$$\mu_{ik}=\left[\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}\right]^{-1},$$

这里取 $m=2$，计算结果列为 μ 矩阵。

$$\mu^{(1)}=\begin{bmatrix}0.986 & 0.979 & 0.438 & 0.019 & 0.013 & 0.062\\ 0.013 & 0.021 & 0.562 & 0.980 & 0.987 & 0.938\end{bmatrix}$$

4° 检 查 结 果：$\max\{|\mu_{ij}-u_{ij}^{0}|\}$ 是否小于 ε，这里取 $\varepsilon=0.1$。不难看出，当 $i=3$，$j=1$，2 时 $\max\{|0.438-0|\}<0.1$ 就不成立。因此需要返回到 2° 重新计算 V_i。

2°¹　以 μ_1 和 V_i 的公式计算 V_i 矩阵：

$$V_1^{(1)}=\begin{bmatrix}1.82\\ 1.27\end{bmatrix},\quad V_2^{(2)}=\begin{bmatrix}8.66\\ 7.56\end{bmatrix}.$$

其中 $\nu_{11}=[(0.986)^2+2(0.979)^2+5(0.438)^2$

$+7(0.019)^2+9(0.013)^2+10(0.062)^2]$

$\div[(0.986)^2+(0.979)^2+(0.438)^2$

$+(0.019)^2+(0.013)^2+(0.062)^2]$

$=3.890786\div2.126855=1.829$

同理算出其它的 ν_{12}，ν_{21}，ν_{22}。

3°¹　计算新的划分矩阵 $\mu^{(2)}$：用(7.3.6)可得距离矩阵 $D=[d_{ik}]$：

$$D=\begin{bmatrix}1.51 & 0.75 & 4.19 & 8.49 & 9.19 & 11.25\\ 10.76 & 8.68 & 5.11 & 1.72 & 0.66 & 1.97\end{bmatrix}$$

于是将 d_{ik} 代入 μ_{ik} 的计算公式，得新的划分矩阵 $\mu^{(2)}$：

$$\mu^{(2)}=\begin{bmatrix}0.98 & 0.99 & 0.59 & 0.04 & 0.01 & 0.03\\ 0.02 & 0.01 & 0.41 & 0.96 & 0.99 & 0.97\end{bmatrix}.$$

4°¹　检查 $\max\{|\mu_{ij}^{(2)}-\mu_{ij}^{(1)}|\}$ 是否小于 0.1.事实上，

$|\mu_{13}^{(2)} - \mu_{13}^{(1)}| = 0.15 > 0.1$，故仍需返回第二步，重新计算聚类中心。下面重复以上手续。

2^{02}　重新计算聚类中心，$V_1^{(2)}$，$V_2^{(2)}$：

$$V_1^{(2)} = \begin{bmatrix} 2.04 \\ 1.47 \end{bmatrix}, \quad V_2^{(2)} = \begin{bmatrix} 4.88 \\ 7.76 \end{bmatrix}$$

3^{02}。计算相应的划分矩阵。为此计算出距离矩阵：

d_{ik}	x_1	x_2	x_3	x_4	x_5	x_6
$V_1^{(1)}$	1.47	0.53	3.89	5.58	8.89	10.96
$V_2^{(2)}$	7.48	8.67	5.12	1.5	0.92	1.96

从而得软划分矩阵 $\mu^{(3)}$ 如下：

$$\mu^{(3)} = \begin{bmatrix} 0.96 & 0.99 & 0.63 & 0.07 & 0.01 & 0.03 \\ 0.04 & 0.01 & 0.37 & 0.93 & 0.99 & 0.97 \end{bmatrix}$$

4^{02}。对于给定的 $\varepsilon = 0.1$，有 $\max\{|\mu_{ij}^{(3)} - \mu_{ij}^{(2)}|\} < 0.1$，（$i = 1, 2$；$j = 1, 2, \cdots, 6$。从而得到 $c = 2$ 时的最佳分类：

$$\tilde{G}_1 = \{0.96/x_1,\ 0.99/x_2,\ 0.63/x_3,\ 0.07/x_4,\ 0.01/x_5,\ 0.03/x_5\};$$

$$\tilde{G}_2 = \{0.04/x_1,\ 0.01/x_2,\ 0.37/x_3,\ 0.93/x_4,\ 0.99/x_5,\ 0.97/x_6\}.$$

于是分类结束。相应的硬分类矩阵为

$$U = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

故

$$G_1 = \{x_1,\ x_2,\ x_3\},$$

$$G_2 = \{x_4,\ x_5,\ x_6\}.$$

通过这个例子，我们可以看出软划分这一聚类法的计算

步骤，并且计算中的循环次数与选定的精度 ε 有很大关系，ε 小则循环次数多。在给定 ε 的情况下，若初始矩阵选择得好，如该例中就将 x_1，x_2，x_3 分为一类，x_4，x_5，x_6 为另一类，则很快就会得出软划分阵，从而得到最终软分类。

§4 不分明 PFS 聚类法

从第三节介绍的硬划分或软划分均可看出，聚类过程工作量是相当大的。1979年 Vogel 和 Wong 提出 *PFS* 聚类法，1985 年张伟提出不分明 *PFS* 聚类法([36][38])。现在介绍这些方法。

7.4.1 *PFS* 聚类法

PFS 一词是"伪 F 统计"的英文 Pseudo F Statistic 的缩写。

设样本集 $X=\{x_1, x_2, \cdots, x_n\}$，$n$ 为有限数，每个样本仅测量了一个指标，即该变量是一维的。现在将其分为 c 个类：G_1，G_2，…，G_c。记 $\boldsymbol{\nu}_i$ 为第 i 类(G_i)的中心。从而样本 x_k 与 $\boldsymbol{\nu}_i$ 的距离平方和为

$$S_W^1=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}(x_k-\nu_i)^2 \qquad (7.4.1)$$

它反映了所有样本到它所属类的聚类中心的距离平方和，为类内样本散布矩阵的表示式。同样考虑该变量样本的类间、混合散布矩阵的表示式分别为

$$S_B^1=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}\nu_i^2, \qquad (7.4.2)$$

$$S_T^1=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}x_k^2. \qquad (7.4.3)$$

其中 μ_{ik} 的取值由下式决定。

$$\mu_{ik}=\begin{cases}0, & x_k \notin G_i, \\ 1, & x_k \in G_i\end{cases} \tag{7.4.4}$$

有了这些准备，我们可以证明下式成立：

$$t_r(S_T^1)=t_r(S_B^1)+t_r(S_W^1) \tag{7.4.5}$$

其中$t_r(\cdot)$表示矩阵·之迹，即矩阵·主对角线元素之和。

由此，对于该变量样本可定义 F 统计比率如下：

$$FS=\frac{t_r(S_B^1)(n-c)}{t_r(S_W^1)(c-1)} \tag{7.4.6}$$

如果样本的变量不是一维的而是 p 维的，即

$$X=\{x_1,x_2,\cdots x_n\} \tag{7.4.7}$$

$$x_k=(x_{k1},x_{k2},\cdots,x_{kp})', k=1,2,\cdots,n.$$

其中“ ′ ”表示转置。现要分为 c 类，相应的聚类中心为 V_i，$i=1, 2, \cdots, c$：

$$V_i=(v_{i1},v_{i2},\cdots,v_{ip})', \tag{7.4.8}$$

象一维情形那样，引入类间、类内及混合散布矩阵S_B^p，S_W^p及S_T^p，其表示式为

$$S_B^p=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}V_iV_i' \tag{7.4.9}$$

$$S_W^p=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}(x_k-V_i)(x_k-V_i)' \tag{7.4.10}$$

$$S_T^p=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}x_kx_k' \tag{7.4.11}$$

仿 FS 的定义，对于 p 维情形定义伪 F 统计比为

$$PFS=\frac{t_r(S_B^*)(n-c)}{t_r(S_W^*)(c-1)}. \tag{7.4.12}$$

至此，我们引出了 PFS。

不难证明，对任意维数的变量样本总有

$$t_r(S_T)=t_r(S_W)+t_r(S_B) \tag{7.4.13}$$

另外，从 PFS 的定义可以看出，当 c 固定时，若要 PFS 最大，必须 $t_r(S_W^*)$为最小。我们知道，对于硬划分或软划分来说，若能使 $t_r(^*_W)$ 达最小，则对应的分类结果，一般说来是人令满意的。

当 c 由小变大时，PFS 值可能在某一 c 值取得最大值。这正是我们所要寻找的最佳分类数。因为 PFS 值大，意味着各个类自身很紧凑，但类间却是清晰可分的。

因此，如果分类结果具有这些性质，那么此结果一定是合理的。

7.4.2 不分明PFS聚类法

由于不分明聚类问题中，已无明确的类内、类间概念，因此象硬划分与软划分的区别那样，将 PFS 聚类法过渡到不分明情形。对于 p 维变量样本集 X及聚类中心同式(7)(8)表示的一样。仿照(9)–(11)式，引入不分明类间、类内及不分明混合散布矩阵 S_{BF}，S_{WF}^* 及 S_{TF}^*，其表示式如下：

$$S_{BF}^*=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}^m V_i V_i' \tag{7.4.14}$$

$$S_{WF}^*=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}^m (x_k-V_i)(x_k-V_i)' \tag{7.4.15}$$

$$S_{TF}^*=\sum_{i=1}^{c}\sum_{k=1}^{n}\mu_{ik}^m x_k x_k' \tag{7.4.16}$$

其中的 m，取值范围为[1,∞)，其作用是为了加强样本对于各类的隶属对比度。而 μ_{ik} 是第 k 个样本隶属于第 i 类的隶属度，取值于[0,1]。

不难证明，S_{BF}^{p}，S_{WF}^{p} 及 S_{TF}^{p} 之间仍存在如下的关系：

$$\mathrm{tr}(S_{TF}^{p}) = \mathrm{tr}(S_{BF}^{p}) + \mathrm{tr}(S_{WF}^{p}).$$

类似地，我们定义"不分明伪 F 统计比率"如下：

$$F\text{-}PFS = \frac{\mathrm{tr}(S_{BF}^{p})(n-c)}{\mathrm{tr}(S_{WF}^{p})(c-1)}.$$

当 c 取不同的值时，$F\text{-}PFS$ 值也将不同。且在 c 取某个值 c_{opt} 时，$F\text{-}PFS$ 达到最大值。此时 c_{opt} 即为最佳分类数。但要求得 c_{opt}，则仍需就 $2 \leqslant c \leqslant n-1$ 值计算 $F\text{-}PFS$ 值，从中选最大值相应之 c_{opt}。

关于 V_i 的选择，文中认为将 Bezdck 提出的软划分中所选之 V_i 及 μ_{ik} 移植过来。即

$$V_i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m x^k}{\sum_{k=1}^{n} (\mu_{ik})^m}, \quad i = 1, 2, \cdots, c,$$

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}},$$

其中 $d_{ik} = \|x_k - V_k\|$，$i = 1, \cdots, c$；$k = 1, 2, \cdots, n$.

在此 $d_{ik} = \left[\sum_{j=1}^{p} (x_{kj} - v_{ij})^2\right]^{1/2}$.

有了以上的介绍，我们不难进行对 X 聚类了。下面给出计算的具体步骤：

1° 给出样本集 $X=\{x_1, x_2, \cdots, x_n\}$，且

$$x_k=(x_{k1}, x_{k2}, \cdots, x_{kp}),\ k=1,2,\cdots,n.$$

2° 给出 c 的取值范围：$2\leqslant c\leqslant n-1$。

3° 给出初始分类方案 CLMO。它可以由软划分得出，亦可人为给定。

4° 计算 V_i，$i=1, 2, \cdots, c$，

5° 计算 $t_r(S^*_{BF})$及 $t_r(S^*_{WF})$之值。

6° 计算 d_{ik} 及μ_{ik}，$i=1, 2, \cdots, c$；$k=1, 2, \cdots, n$。得到新的分类方案 CLM。

7° 比较 CLM 和 $CLMO$。如果$\|CLM-CLMO\|\leqslant\varepsilon$（$\varepsilon$为事先给定的一个小正数），则往下执行，否则返回到 4°。

8° 计算 F-PFS 之值。

9° 令 c 增加到 $c+1$，如果 $c\leqslant c_{max}$，则返回到 3°，否则停止。

以上步骤可以用计算机进行。文[36]给出其框图及一个例子，说明方法的有效性。

设样本集 $X=\{x_1, x_2, \cdots, x_{15}\}\in R^4$，具体数据列于表 1（表中数据扩大了 100 倍）：

表 1

P_i \ x_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	95	93	91	92	85	86	83	81	20	18	22	19	35	30	32
2	54	58	61	57	48	50	51	49	60	57	61	56	55	52	54
3	38	40	42	43	35	36	38	37	−10	−13	−15	−16	10	08	05
4	20	25	27	23	28	22	24	29	55	53	51	56	50	56	60

计算中取 c 的最小值为 2，最大值为 8，初始分类通过

不分明系统聚类法获得。表 2 列出了各个 c 值所对应的 F-PFS 之值.

可见当 $c=4$ 时，F-PFS 为最大。故 $c=4$为我们所寻求

表 2

C	2	3	4	5	6	7	8
F-PFS	1150	1446	1923	1719	1576	1473	1315

的最佳分类数。此时对应的不分明分类结果列于表 3。

表 3

μ_{ik}	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
G_1	.800	.986	.891	.964	.052	.133	.054	.069
G_2	.190	.013	.102	.034	.944	.862	.943	.924
G_3	.004	.000	.003	.001	.001	.002	.001	.003
G_4	.006	.000	.004	.001	.002	.003	.002	.004

μ_{ik}	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
G_1	.002	.001	.002	.002	.008	.001	.004
G_2	.002	.001	.003	.002	.011	.002	.006
G_3	.905	.989	.965	.972	.051	.013	.048
G_4	.031	.009	.030	.024	.930	.984	.942

取每列中的最大隶属度为 1，其余均为 0，则得相应的硬划分结果。它把上面的 15 个样本分成如下四类：

G_1: $\{x_1, x_2, x_3, x_4\}$,

G_2: $\{x_5, x_6, x_7, x_8\}$,

G_3: $\{x_9, x_{10}, x_{11}, x_{12}\}$,

G_4: $\{x_{13}, x_{14}, x_{15}\}$.

凭直观就可知道，该分类结果是十分合理的。

§5 补 记

设样本集 $X=\{x_1, \cdots, x_n\}$，现将其按照某种意义下的关系分为几族 $G_1, G_2, \cdots, G_c$，使得

$$\bigcup_{i=1}^{c} G_i = X, \quad G_i \cap G_j = \varnothing, \quad i \neq j,$$

$$i, \ j = 1, \ \cdots, \ c.$$

这种族称为类，而把 X 分为类称为聚类。这里事先没有什么可供参考或依循，要说有，只能是样本之间或其特性之间的某种关系，但对 G_i 却一无所知。如何合理的分类，这就是聚类分析。关于通常的聚类分析，已如前述，内容十分广泛，而利用不分明关系进行聚类不过是近十多年的事。1970年有人给出用 c 个不分明集合来划分 X 的方法，引入表示聚类好坏的评价函数，从而找出使该函数达最小的这 c 个不分明集。1974年 Dunn 及 Bezdek[39,40]解决了不分明分类的最优化问题。而聚类的有效性问题，Bezdek(1975)，Widham(1981)，LIU LaFu(1984)又进一步作了研究[41,42]。

关于 X 上一个不分明集 A，如何对它们进行分类的问题，也有多人对此进行研讨。这种问题往往是将 A 中所含的样本分到已知的确定的几个类中。按一般认识，它属于模式

识别问题，这里就从略了，可见[43]。

聚类分析与模式识别之间是有差别的，但彼此之间又有所联系，虽然理论方面尚未完备，却有着广泛的应用和广阔的前景。

参 考 文 献

[1] 张文修 模糊数学基础 西安交大出版社 1984.

[2] 汪培庄 模糊集合论及其应用 上海科技出版社 1984.

[3] Kaufmann A., Introduction to the Theory of Fuzzy Subsets, Vol.4.

[4] 浅居喜代治等 赵汝怀译 模糊系统理论入门 北京师范大学出版社 1982.

[5] 贺仲雄 模糊数学及其应用 天津科学技术出版社 1982.

[6] 郑兴国 模糊性、随机性与可靠性教学与研究 第二炮兵学院 1985.3.

[7] Jain R., Fuzzism and Real world Problems, Fuzzy Sets, P.129-132.

[8] Zadeh L. A., Fuzzy Sets. Inf. and Control 8(1965),338-353.

[9] Zadeh L. A., Probability of Fuzzy Events, J.Math. Anal.Appl., 23(1968), 421-427

[10] 朱永庚 关于模糊性度量公理的注记 模糊数学 1986.第1期

[11] 潘雪海 不分明事件的概率测度 计算机应用与应用数学 1978.第8期

[12] 何家儒 不分明事件与不分明概率 四川师院学报(自然科学)

[13] 仲崇骥 Fuzzy 概率空间 吉林师大学报（自然科学）

1978. 第 1 期

[14] 陈绍仲 刘作述 Fuzzy 概率的一些问题 模糊数学 1982.第 4 期

[15] Klement E.P., Fuzzy σ-Algebra and Fuzzy Measurable Function. Fuzzy Sets and Systems.(1980)83-93

[16] Klement E.P., Lowen R., Schwyhla w., Fuzzy Pro. Measure. Fuzzy Sets and Systems. 5(1981)21-30.

[17] Smets P. Pro. of a Fuzzy Event: and axiomatic approach. Fuzzy Sets and Systems. 7(1982) 153-164

[18] Yager R.R., Probabilities from Fuzzy Observations. Inf. Sciences, 32(1984) 1-31.

[19] Stein W.E. and Talati K., Covex Fuzzy Random Variables. Fuzzy Sets and System, 6(1981) 271-283

[20] Kwakernaak H., Fuzzy Random Variable-Ⅰ Definitions and Theorems. Inf. Sci. 15(1978) 1-29.

[21] Kwakernaak H., Fuzzy Random Variables-Ⅱ. Algorithems and Examples for the Discrete Case. Inf. Sci., 17(1979)253-278

[22] Nahmias S., Discrete Fuzzy Random Variables. Technical Report 18

[23] Nahmias S., Fuzzy Variables. Fuzzy Sets and Systems. 1(1978) 97-110.

[24] Piasecki K., Probability of Fuzzy Events Defined as Denumerable Additivity Measure, Fuzzy Sets and Systems 17(1985), 271-284.

[25] Piasecki K., Extension of Fuzzy P-measure, BUSEFAL 19(1984)26-41

[26] Piasecki K., On One Relationship between Chas-

sical Probability Measure and Fuzzy P-Measure, BUSEFAL 24(1985), 29-41.

[27] Piasecki K., On the Bayes Formula for Fuzzy Probability Measures, Fuzzy Sets and Systems. 18(1986), 183-185

[28] Piaseck K.,Extension of Fuzzy P-Measure Generated by usual Measure, 模糊数学 3, 4(1987)117-124.

[29] Sugeno M., Theory of Fuzzy Intergrals and Its Applications, ph. D. Thesis, Tokyo. Inst. of Technol 1970.

[30] Dubois D. and Prade H., Fuzzy Sets and Systems (Theory and Applications),Academic Press, 1980.

[31] 张尧庭 多元统计分析引论 科学出版社 1982.

[32] 张南纶 随机现象的从属特性与概率特性 武汉建材学院学报 1981.№.1-2.

[33] 袁嘉祖 钱林清 Fuzzy $L_9(3^4)$正交设计 模糊数学 1986.12. 25-30.

[34] 汪贤裕 Fuzzy 线性拟合 成都科技大学学报 1987.№4.

[35] 北大数力系 正交设计 人民出版社

[36] 张 伟 Fuzzy 聚类法中的一个新算法：Fuzzy PFS 聚类法 模糊数学 1987. № 3-4. 51-55.

[37] Diday E., (1974), Recent Progress in Distance and Similarity Measures in Pattern Recognition, Second International Joint Conference on Pattern Recognition, 1535-1939.

[38] Vogel M.A.,Wong A.C.,PFS Clustering Method. IEEE Trans, Vol. PAMI-1, №.3, 1979.

[39] J.C. rical taxonomy wtih
Fuzzy cal Biology, 1, 57-71.
[40] J.C. Dunn, A Fu lative of the ISODATA Process and its use in Detecting well-Separated Clusters,J. of Cybernetics, 3. 32-47, (1974).
[41] Widham, M. P., (1981) Cluster Vilidity for Fuzzy Clustering Algerithm, Fuzzy Sets and Systems,5, 177-185.
[42] Bezdek, J.C., Pattern Rencognition with Fuzzy Objective Function Algorithm, Plenum Press, New York, 1981.
[43] 于纯海 Fuzzy 集上的 Fuzzy 关系及其分类 模糊数学 1987. №3-4.
[44] 赵德玉 模糊聚类分析在啤酒大麦育种上的应用 模糊数学 1987.1.
[45] 虞英军 Fuzzy 聚类分析在零件分类编组中的应用 模糊数学 1987.2.